Regression III

Linear Model Interactions

Dave Armstrong

Goals for Today

1. Interaction effects for 2 categorical variables
2. Interaction effects for categorical and quantitative variables.
    - Dummy-quantitative interaction
    - Categorical-quantitative interaction
3. Interaction effects for 2 quantitative variables

Interaction Effects (1)

When the partial effect of one variable depends on the value of another variable, those two variables are said to "interact".

- For example, we may want to test whether age effects are different for men (coded 1) and women (coded 0).
- In such cases it is sensible to fit separate regressions for men and women, but this does not allow for a formal statistical test of the differences
- Specification of interaction effects facilitates statistical tests for a difference in slopes within a single regression

Notes

Type notes here...

Interaction Effects (2)

Interaction terms are the *product of the regressors for the two variables*.

- The interaction regressor in the model below is X_iD_i:

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta(X_iD_i) + \varepsilon_i$$
$$\text{income}_i = \alpha + \beta\,\text{age}_i + \gamma\,\text{men}_i + \delta(\text{age}_i \times \text{men}_i) + \varepsilon_i$$

Ultimately we want to know two things:

- Is there a statistically significant interactive (i.e., multiplicative or conditional) effect?
- If the answer to #1 is "yes", what is the nature of that effect (i.e., what does it look like)?

Below, I will walk you through all of the possible two-way interaction scenarios and we will discuss how to answer these two questions.

Notes

Type notes here...

ANOVA Type I Sums of Squares

Consider the model:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1x_2 + e$$

In a type I test, the following tests are calculated.

1. The effect of x_1 not controlling for any other variables.
2. The effect of x_2 controlling for x_1.
3. The effect of x_3 controlling for x_1 and x_2.
4. The effect of the interaction, x_1x_2 controlling for x_1, x_2 and x_3.

The results depend on the order in which the variables are included in the model.

The `anova()` function in the `stats` package does this kind of test.

Notes

Type notes here...

ANOVA Type II Sums of Squares

Consider the model:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1x_2 + e$$

In a type II test, the following tests are calculated.

1. The effect of x_1 controlling for x_2 and x_3.
2. The effect of x_2 controlling for x_1 and x_3.
3. The effect of x_3 controlling for x_1 and x_2 and x_1x_2.
4. The effect of the interaction, x_1x_2 controlling for x_1, x_2 and x_3.

When testing lower-order terms, they do not control for higher-order terms of the same variable(s).

The `ANOVA(..., type="II")` function in the `car` package does this test.

Notes

Type notes here...

ANOVA Type III Sums of Squares

Consider the model:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_1 x_2 + e$$

In a type III test, the following tests are calculated.

1. The effect of x_1 controlling for x_2, $x_1 x_2$ and x_3.
2. The effect of x_2 controlling for x_1, $x_1 x_2$ and x_3.
3. The effect of x_3 controlling for x_1, x_2 and $x_1 x_2$.
4. The effect of the interaction, $x_1 x_2$ controlling for x_1, x_2 and x_3.

When testing lower-order terms, they do control for higher-order terms of the same variable(s).

The `ANOVA(..., type="III")` function in the `car` package does this test.

Notes

Type notes here...

Two Categorical Variables

With two categorical variables, essentially you are estimating a different conditional mean for every pair of values across the two categorical variables. You could do that as follows:

```
library(DAMisc)
library(car)
data(Duncan)
Duncan <- Duncan %>%
  mutate(inc.cat = cut(Duncan$income, 3),
         inc.cat = factor(as.numeric(inc.cat),
                          labels=c("Low", "Middl

mod <- lm(prestige~ inc.cat * type + education,
  data=Duncan)
```

```
S(mod, brief=TRUE)
```

```
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)              7.8827     3.4364   2.294 0.027915 *
## inc.catMiddle           22.4574     4.8792   4.603 5.30e-05 ***
## inc.catHigh             51.2807     9.4351   5.435 4.29e-06 ***
## typeprof                55.6073    11.6800   4.761 3.30e-05 ***
## typewc                   2.5446     8.1162   0.314 0.755746
## education                0.2799     0.1121   2.496 0.017411 *
## inc.catMiddle:typeprof -41.5789    11.2428  -3.698 0.000740 ***
## inc.catHigh:typeprof   -50.3567    13.3929  -3.760 0.000621 ***
## inc.catMiddle:typewc   -13.0171    10.3130  -1.262 0.215223
## inc.catHigh:typewc     -33.6407    13.1215  -2.564 0.014806 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard deviation: 9.115 on 35 degrees of freedom
## Multiple R-squared: 0.9334
## F-statistic: 54.54 on 9 and 35 DF,  p-value: < 2.2e-16
##    AIC    BIC
## 337.29 357.16
```

Notes

Type notes here...

Anova

Q1: Is there an interaction Effect here?

- An incremental (Type II) F-test will answer that question. We want to test the null hypothesis that all of the interaction dummy regressor coefficients are zero in the population.
- The `inc.cat:type` line of the output gives the results of this test.

```
Anova(mod)
```

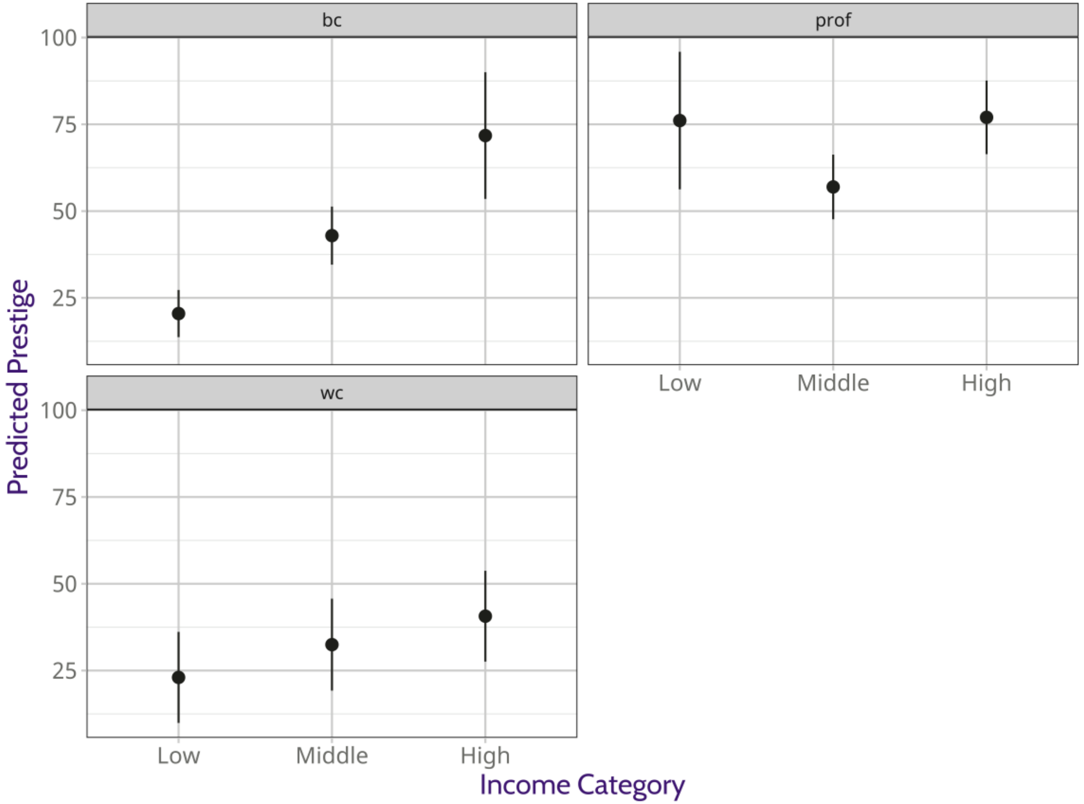
```
## Anova Table (Type II tests)
##
## Response: prestige
##              Sum Sq Df F value    Pr(>F)
## inc.cat      3491.9  2 21.0159 1.010e-06 ***
## type         2856.0  2 17.1885 6.308e-06 ***
## education     517.7  1  6.2313  0.017411 *
## inc.cat:type 1644.4  4  4.9484  0.002871 **
## Residuals    2907.7 35
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Notes

Type notes here...

Q2: What is the nature of the interaction?

```r
library(ggeffects)
e1 <- ggpredict(mod,
                terms=c("inc.cat", "type"))
ggplot(e1) +
  geom_pointrange(aes(x=x, y=predicted,
                      ymin=conf.low,
                      ymax=conf.high)) +
  facet_wrap(~group, ncol=2) +
  theme_bw() +
  mytheme() +
  labs(x="Income Category",
       y="Predicted Prestige")
```

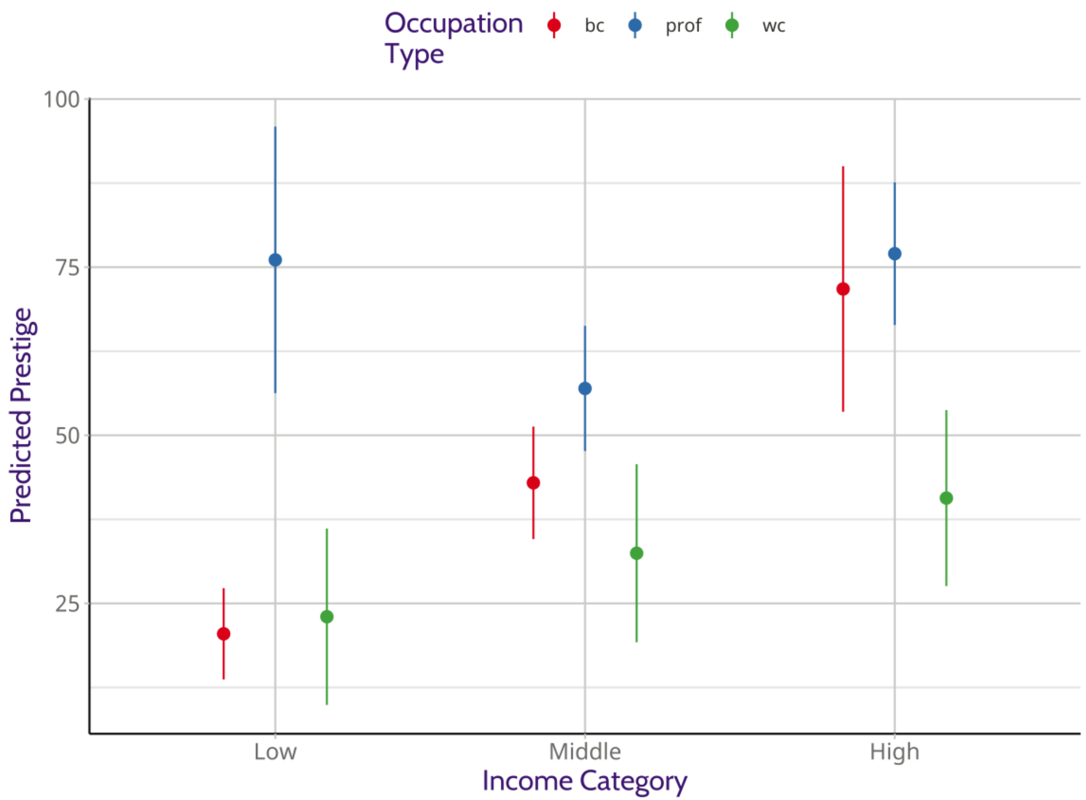


Notes

Type notes here...

Or Alternatively

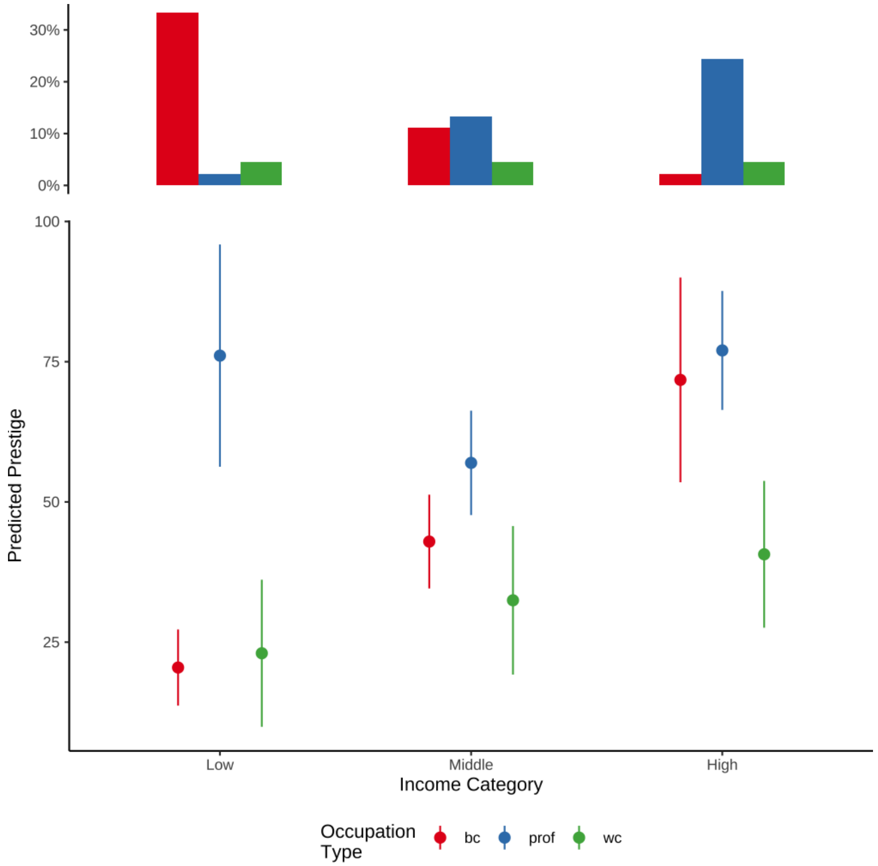
```
ggplot(e1) +
  geom_pointrange(aes(x=x, y=predicted,
                      ymin=conf.low,
                      ymax=conf.high,
                      colour=group),
                      position = position_dodge(width
  theme_classic() +
  scale_color_brewer(palette="Set1") +
  mytheme(legend.position="top") +
  labs(x="Income Category",
       y="Predicted Prestige",
       colour="Occupation\nType")
```



Notes

Type notes here...

With Bar Density



Notes

Type notes here...

Interpretation

The important points are as follows:

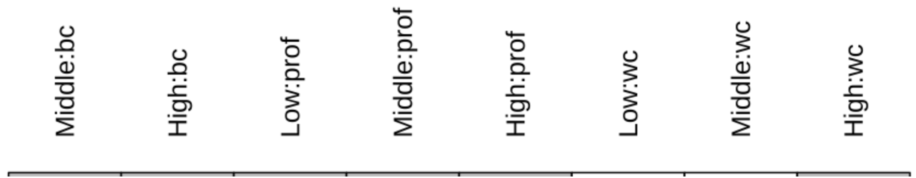
- The interaction term is significant in the F-test, so that indicates a significant interaction effect.
- With no interaction effect, the across each row have the same pattern across the three different tows and down the three different columns.
- While the trends overall look somewhat different and there are clearly different magnitudes in the differences.
- This is the same as we look down the rows.

Notes

Type notes here...

Using Factorplot

```
library(factorplot)
library(effects)
e <- effect("inc.cat*type", mod)
fp <- factorplot(e)
plot(fp, print.square.leg=F,
     scale.text=.75, abbrev.char=100)
```



Notes

Type notes here...

Testing Differences

Imagine that you wanted to test whether the effect of moving from middle income to high income was the same for blue collar and white collar occupations.

$$\begin{aligned}\hat{P} &= b_0 + b_1 M + b_2 H + b_3 Pr + b_4 W + b_5 E \\ &\quad + b_6 M \times Pr + b_7 H \times Pr + b_8 M \times W + b_9 H \times W\end{aligned}$$

The effect for blue collar occupations is:

$$b_2 - b_1$$

And for white collar occupations it is

$$(b_2 + b_9) - (b_1 + b_8)$$

Notes

Type notes here...

Rearranging, we get:

$$
\begin{aligned}
b_2 - b_1 &= (b_2 + b_9) - (b_1 + b_8) \\
&= b_2 + b_9 - b_1 - b_8 \\
0 &= b_9 - b_8
\end{aligned}
$$

```
linearHypothesis(mod,
                "inc.catHigh:typewc - inc.catMiddle:typewc = 0")
```

```
## Linear hypothesis test
##
## Hypothesis:
## - inc.catMiddle:typewc  + inc.catHigh:typewc = 0
##
## Model 1: restricted model
## Model 2: prestige ~ inc.cat * type + education
##
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     36 3100.9
## 2     35 2907.7  1    193.19 2.3254 0.1363
```

Notes

Type notes here...

Two Non-Reference Categories

What if we want to test whether the effect of middle to high income is different for Professional and White Collar occupations? The effect for Professional Occupations is:

$$(b_2 + b_7) - (b_1 + b_6)$$

Thus, the difference in effects is:

$$\begin{aligned} b_2 + b_7 - b_1 - b_6 &= b_2 + b_9 - b_1 - b_8 \\ b_7 - b_6 &= b_9 - b_8 \\ 0 &= b_6 - b_7 + b_9 - b_8 \end{aligned}$$

Notes

Type notes here...

The test

```
linearHypothesis(mod,
                "inc.catMiddle:typeprof -inc.catHigh:typeprof +
                 inc.catHigh:typewc - inc.catMiddle:typewc = 0")
```

```
## Linear hypothesis test
##
## Hypothesis:
## inc.catMiddle:typeprof - inc.catHigh:typeprof - inc.catMiddle:typewc  + inc.catHigh:typewc = 0
##
## Model 1: restricted model
## Model 2: prestige ~ inc.cat * type + education
##
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     36 3015.2
## 2     35 2907.7  1    107.52 1.2942  0.263
```

Notes

Type notes here...

One Dummy and One Continuous

$$Y_i = \alpha + \beta X_i + \gamma D_i + \delta(X_i D_i) + \varepsilon_i$$

One way to think about this model is leading to two separate regression lines:

For D = 0:

$$\begin{aligned} \hat{Y}_i &= \alpha + \beta X_i + \gamma(0) + \delta(X_i \times 0) \\ &= \alpha + \beta X_i \end{aligned}$$

For D=1:

$$\begin{aligned} \hat{Y}_i &= \alpha + \beta X_i + \gamma(1) + \delta(X_i \times 1) \\ &= (\alpha + \gamma) + (\beta + \delta) X_i \end{aligned}$$

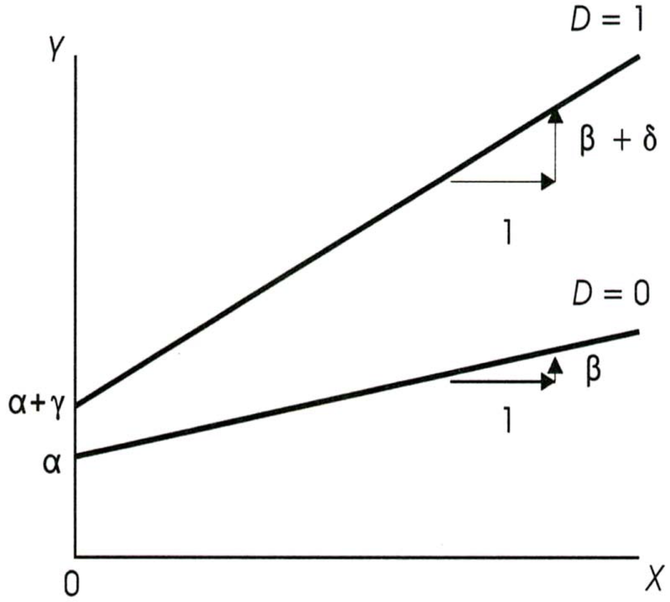


Figure 7.5 from Fox (1997)

Notes

Type notes here...

Example with one Dummy Variable and One Continuous Variable

```r
library(car)
data(SLID)
mod <- lm(wages ~ age*sex, data=SLID)
S(mod, brief=TRUE)
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.84674    0.50267  15.610  < 2e-16 ***
## age          0.16377    0.01295  12.648  < 2e-16 ***
## sexMale     -1.78986    0.70988  -2.521   0.0117 *
## age:sexMale  0.13625    0.01820   7.485 8.71e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard deviation: 7.122 on 4143 degrees of freedom
##   (3278 observations deleted due to missingness)
## Multiple R-squared: 0.1844
## F-statistic: 312.3 on 3 and 4143 DF,  p-value: < 2.2e-16
##      AIC      BIC
## 28057.09 28088.74
```

Notes

Type notes here...

Assessing Interaction I

Q1: Is there an interaction?

- We want to know whether the lines are parallel or not.
- Note that the coefficient on the interaction term gives the difference in the slope for the $D = 0$ group and the $D = 1$ group.
- The `age:sexMale` line provides the answer to the question.

The answer ...

- If the coefficient is statistically significant (and it is here), then there is a significant interaction.
- If the coefficient is not statistically significant, then a purely additive model performs just as well.

Notes

Type notes here...

Q2: What is the nature of the interaction?

There are a number of ways we can figure this out. Ultimately, we want to know three things regarding the slope.

- Is the slope of age for females ($D = 0$) different from zero?
- Is the slope of age for males ($D = 1$) different from zero?
- Is the slope of age for men different from the slope of age for women?

Two of these can be answered directly from the coefficient table, one requires a bit of extra work.

Notes

Type notes here...

Conditional Effect of Age

First, we need to think more generally about the conditional effect of age. If the equation is:

$$\text{wages} = b_0 + b_1\text{age} + b_2\text{male} + b_3\text{age} \times \text{male} + e$$

Then the partial, conditional effect (or what some might call the "marginal effect") of age is:

$$\frac{\partial\widehat{\text{wages}}}{\partial\text{age}} = b_1 + b_3\text{male}$$

Since we will want to test hypotheses about that quantity, we need to know its variance:

$$V(b_1 + b_3\text{male}) = V(b_1) + \text{male}^2 V(b_3) + 2\text{male}V(b_1, b_3)$$

In general, with constants c and d and variables W and Z:

$$V(cW + dZ) = c^2V(W) + d^2V(Z) + 2cdV(W, Z)$$

Notes

Type notes here...

Back to the Questions

- Is the slope of age for females ($D = 0$) different from zero?
  - This amounts to a test of $H_0 : \beta_1 = 0$. This can be evaluated by looking at the `age` line from the output.
- Is the slope of age for men different from the slope of age for women?
  - This amounts to a test of $H_0 : \beta_3 = 0$. This can be evaluated by looking at the `age:sexMale` line from the output.

Notes

Type notes here...

Back to the Questions (2)

- Is the slope of age for males ($D = 1$) different from zero?
  - This amounts to a test of $H_0 : \beta_1 + \beta_3 = 0$. This cannot be directly evaluated by looking at the coefficients. It can be done this way:

```
library(psre)
simple_slopes(mod, "age", "sex")
```

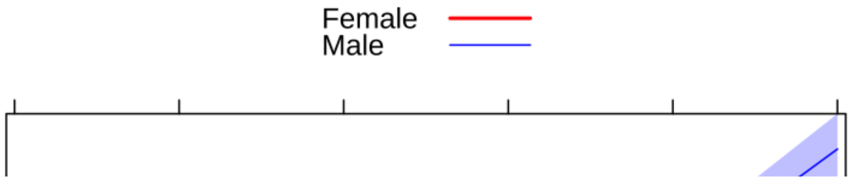
```
## Simple Slopes:
## # A tibble: 2 x 5
##   group  slope     se     t         p
##   <chr>  <dbl>  <dbl> <dbl>     <dbl>
## 1 Female 0.164 0.0129  12.6 5.25e- 36
## 2 Male   0.300 0.0128  23.4 2.89e-114
##
## Pairwise Comparisons:
## # A tibble: 1 x 5
##   comp          diff     se     t         p
##   <chr>        <dbl>  <dbl> <dbl>     <dbl>
## 1 Female-Male -0.136 0.0182 -7.48 8.71e-14
```

Notes

Type notes here...

Graphically...

```r
library(lattice)
trellis.par.set(
  superpose.line=list(col=c("red", "blue")),
  superpose.polygon = list(col=c("red", "blue"))
intQualQuant(mod, c("age", "sex"), type="slopes"
             plot=TRUE, rug=TRUE, ci=TRUE)
```



Notes

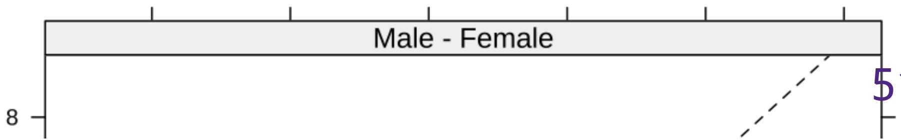
Type notes here...

The effect of Gender

Almost always, we are concerned with the results above (i.e., the different slopes for age), but what if we care about the conditional effect of gender?

$$\frac{\partial \widehat{\text{wages}}}{\partial \text{male}} = b_2 + b_3\text{age}$$

```
intQualQuant(mod, c("age", "sex"),
             type="facs", plot=TRUE)
```



Notes

Type notes here...

Summary

- The interaction is significant (from the `age:sexMale` line of the regression output), so the two variable do have an interactive effect.
- Since the `age` coefficient is positive and the `age:sexMale` coefficient is positive, both men and women have positive slopes of age for wages, but the difference between men and women is significantly bigger than zero, meaning the slope of age for men is bigger than the slope of age for women.
- The results of the `intQualQuant` function (from the `DAMisc` package) provide graphical and numerical results about the two different slopes.
- The above implies that the effect of gender is increasing in age (i.e., the gender gap is growing). The `intQualQuant` function (from the `DAMisc` package) provides numerical and optional graphical results.

Notes

Type notes here...

One Categorical and One Continuous

With one categorical and one continuous variable, we want to show the conditional coefficients of the continuous variable (probably in a table) and we want to show the conditional coefficients of the dummy variables.

```
Prestige$income <- Prestige$income/1000
mod <- lm(prestige ~ income*type + education,
        data=Prestige)
S(mod, brief=TRUE)
```

```
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)    
## (Intercept)       -6.7273     4.9515  -1.359   0.1776    
## income             3.1344     0.5215   6.010 3.79e-08 ***
## typeprof          25.1724     5.4670   4.604 1.34e-05 ***
## typewc             7.1375     5.2898   1.349   0.1806    
## education          3.0397     0.6004   5.063 2.14e-06 ***
## income:typeprof   -2.5102     0.5530  -4.539 1.72e-05 ***
## income:typewc     -1.4856     0.8720  -1.704   0.0919 .  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard deviation: 6.455 on 91 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared: 0.8663
## F-statistic: 98.23 on 6 and 91 DF,  p-value: < 2.2e-16
##    AIC    BIC
```

Notes

Type notes here...

Anova

Q1: Is there a significant interaction?

```
Anova(mod)
```

```
## Anova Table (Type II tests)
##
## Response: prestige
##              Sum Sq Df F value    Pr(>F)
## income       1058.8  1 25.4132 2.342e-06 ***
## type          591.2  2  7.0947   0.00137 **
## education    1068.0  1 25.6344 2.142e-06 ***
## income:type   890.0  2 10.6814 6.809e-05 ***
## Residuals    3791.3 91
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Notice that the `income:type` line of the `Anova` output tells us that the interaction is significant. Thus, we should go on to calculate and explain the conditional coefficients.

Notes

Type notes here...

Conditional Coefficients of Income

Q2: What is the nature of the interaction effect?

- The nature of the interaction has to be considered both for `income` and for `type`.
- We can calculate the conditional effects and variances of `income` as follows:

```r
simple_slopes(mod, "income", "type")
```

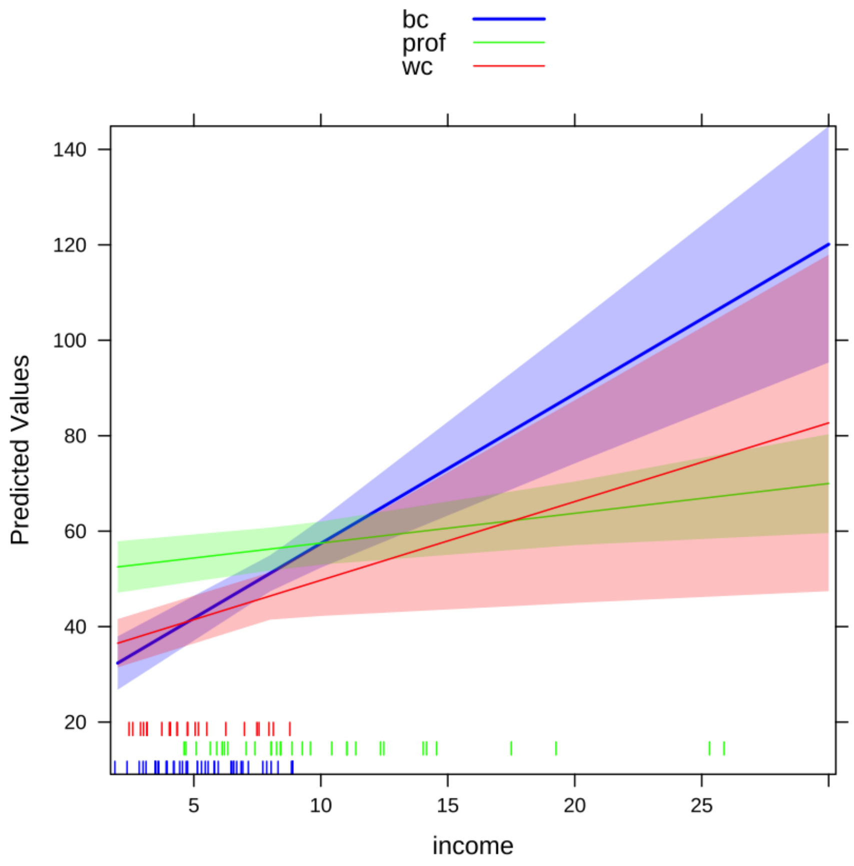
```
## Simple Slopes:
## # A tibble: 3 x 5
##   group slope    se     t            p
##   <chr> <dbl> <dbl> <dbl>        <dbl>
## 1 bc    3.13  0.522  6.01 0.0000000379
## 2 prof  0.624 0.222  2.82 0.00596
## 3 wc    1.65  0.709  2.33 0.0222
##
## Pairwise Comparisons:
## # A tibble: 3 x 5
##   comp     diff    se     t         p
##   <chr>   <dbl> <dbl> <dbl>     <dbl>
## 1 bc-prof  2.51 0.553  4.54 0.0000172
## 2 bc-wc    1.49 0.872  1.70 0.0919
## 3 prof-wc -1.02 0.740 -1.38 0.170
```

Notes

Type notes here...

Conditional Effects of Income

```r
cols <- c("blue", "green", "red")
trellis.par.set(
superpose.line = list(col=cols),
superpose.polygon = list(col=cols))
intQualQuant(mod, c("income", "type"),
             type="slopes", plot=TRUE)
```



Notes

Type notes here...

Interpretation

- The slope is significant for all occupation types and is the biggest for blue collar.
- Confidence bounds for both blue collar and white collar occupation lines are very big at high levels of income (lack of data density).
- The only valid places where professional occupations can be compared to the others is between around 5,000 and 8,000 dollars.

Notes

Type notes here...

Conditional Effect of Type

Q2: What is the nature of the interaction effect (this time for `type`)?

- The conditional effect of type (as we saw) is a bit more difficult. Here, We would presumably have to test each pairwise difference: BC vs Prof, BC vs WC and Prof vs WC for different values of education. First, let's think about what we need.

$$\text{BC vs Prof: } \frac{\partial \text{Prestige}}{\partial \text{Prof}} = b_2 + b_5 \text{Income}$$

$$\text{BC vs WC: } \frac{\partial \text{Prestige}}{\partial \text{WC}} = b_3 + b_6 \text{Income}$$

$$\text{Prof vs WC: } \frac{\partial \text{Prestige}}{\partial \text{Prof}} - \frac{\partial \text{Prestige}}{\partial \text{WC}} = (b_2 - b_3) + (b_5 - b_6) \text{Income}$$

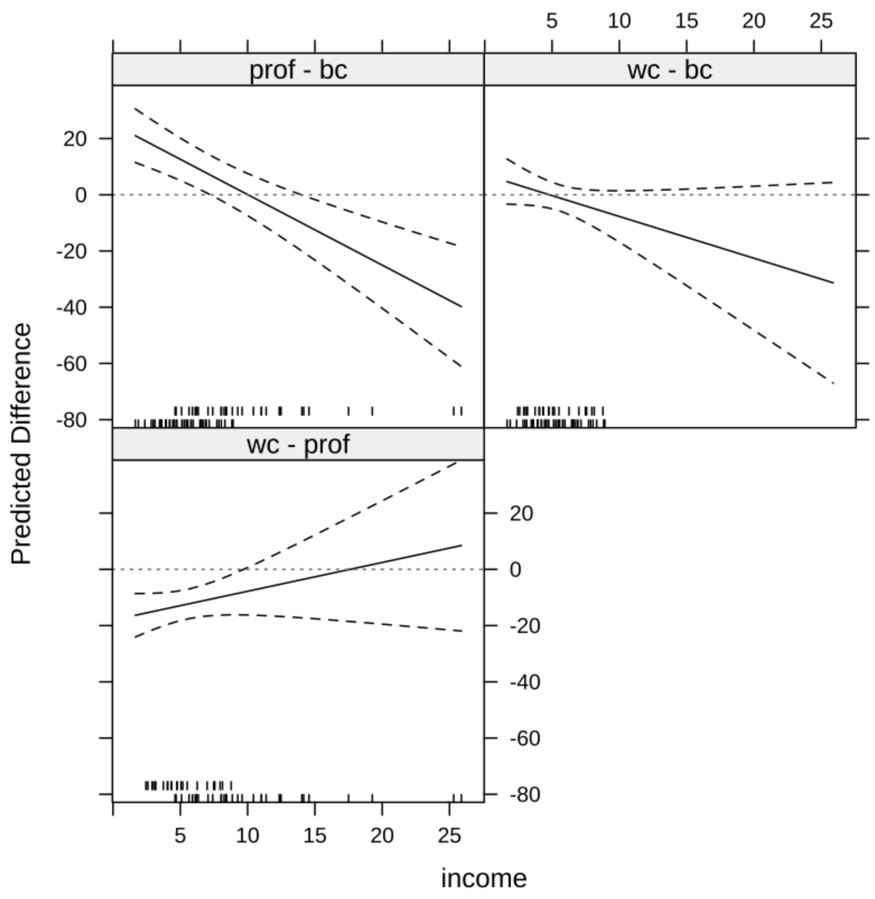
Notes

Type notes here...

Conditional Effect of Type

The conditional effect of type is a bit more difficult, luckily a function exists to help. Here, We would want to test each pairwise difference: BC vs Prof, BC vs WC and Prof vs WC.

```
mod.out <- intQualQuant(mod, c("income", "type"),
  type="facs", n=25, plot=T)
update(mod.out, layout=c(2,2),
       as.table=TRUE)
```



Notes

Type notes here...

Interpretation

In the previous graph, we see the following:

- From its lowest values through the mean of income, professional occupations are expected to have more prestige than blue collar occupations. However, when income is highest, blue collar occupations are expected to have more prestige than professional occupations (first row of table)
- The difference between white collar and blue collar is never significantly different from zero (second row of table).
- From its lowest values through the mean of income, professional occupations are expected to have more prestige than white collar occupations. When income is high, however, there is no expected difference between professional and white collar occupations as regards prestige.

Notes

Type notes here...

Two continuous Variables

With two continuous variables the interpretation gets a bit trickier. For example, consider the following model:

$$\hat{Y}_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i1} X_{i2}$$

We want to know the partial conditional effect of both X_1 and X_2, but unlike above, neither can be boiled down to a small set of values. Just think about the equation:

$$\frac{\partial \hat{Y}}{\partial X_1} = \beta_1 + \beta_4 X_2$$
$$\frac{\partial \hat{Y}}{\partial X_2} = \beta_2 + \beta_4 X_1$$

Note, that β_4 is the amount by which the *effect* of X_1 goes up for every additional unit of X_2 and the amount by which the *effect* of X_2 goes up for every additional unit of X_1.

Notes

Type notes here...

Variance of a Linear Combination

Ultimately, we will want to know when conditional effects are significantly different from zero. This requires us to be able to calculate the variance of the conditional effects.

- Since these are linear combinations of random variables - $\widehat{\beta}_1$, $\widehat{\beta}_2$, and $\widehat{\beta}_4$ and the constants X_1 and X_2, its variance can be easily calculated.

The results above are useful, but these terms get complicated to calculate "by hand" if there is are more than 2 terms for which you want to calculate the variance.

Notes

Type notes here...

Variance of Conditional Effects in Matrix Form

The variance is the sum of all the variance and 2 times all of the pairwise covariances

$$\mathbf{A} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \quad V(\mathbf{W}) = \begin{bmatrix} V(w_1) & V(w_1, w_2) & \cdots & V(w_1, w_k) \\ V(w_2, w_1) & V(w_2) & \cdots & V(w_2, w_k) \\ \vdots & \vdots & \ddots & \vdots \\ V(w_k, w_1) & V(w_k, w_2) & \cdots & V(w_k) \end{bmatrix}$$

Then,

$$V(\mathbf{A}'\mathbf{W}) = \mathbf{A}'V(\mathbf{W})\mathbf{A}$$

Notes

Type notes here...

Testable Hypotheses

$$\hat{Y}_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i1} X_{i2}$$

Berry, Golder and Milton (2012) suggest that we should be able to test 5 hypotheses:

- $\mathbf{P}_{X_1|X_2=\text{min}}$ The marginal effect of X_1 is [positive, zero, negative] when X_2 takes its lowest value.
- $\mathbf{P}_{X_1|X_2=\text{max}}$ The marginal effect of X_1 is [positive, zero, negative] when X_2 takes its highest value.
- $\mathbf{P}_{X_2|X_1=\text{min}}$ The marginal effect of X_2 is [positive, zero, negative] when X_1 takes its lowest value.
- $\mathbf{P}_{X_2|X_1=\text{max}}$ The marginal effect of X_2 is [positive, zero, negative] when X_1 takes its highest value.
- $\mathbf{P}_{X_1 X_2}$ The marginal effect of each of X_1 and X_2 is [positively, negatively] related to the other variable.

Notes

Type notes here...

Example

```
mod <- lm(prestige ~ income*education + type, data=Prestige)
S(mod, brief=TRUE)
```

```
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -17.80359    7.59424  -2.344 0.021212 *
## income              3.78593    0.94453   4.008 0.000124 ***
## education           5.10432    0.77665   6.572 2.93e-09 ***
## typeprof            5.47866    3.71385   1.475 0.143574
## typewc             -3.58387    2.42775  -1.476 0.143303
## income:education   -0.21019    0.06977  -3.012 0.003347 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard deviation: 6.806 on 92 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared: 0.8497
## F-statistic:   104 on 5 and 92 DF,  p-value: < 2.2e-16
##     AIC    BIC
## 661.80 679.89
```

Notes

Type notes here...

Example (2)

Q1: Is there a significant interaction?

- The `income:education` line answers this question. If it is significant, then there is a significant interaction, otherwise there is not.
- This is counter to a minor, though still influential, point in Brambor, Clark and Golder (2006), but is consistent with Berry, Golder and Milton (2012).
- In this case, the interaction is significant, so we can move on to the next question

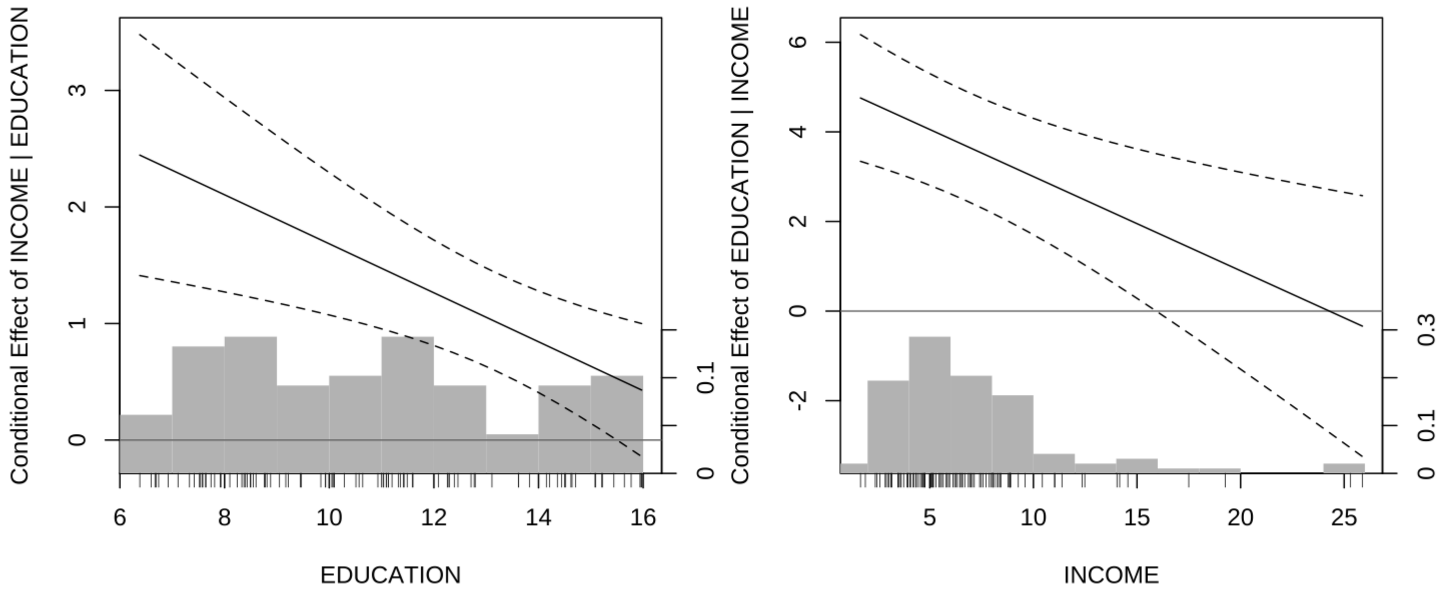
Notes

Type notes here...

Q2: What is the nature of the interaction?

This needs to be shown visually, since there are an infinite number of possibilities.

```
DAintfun2(mod, c("income", "education"), hist=T,
        scale.hist=.3)
```



Notes

Type notes here...

Interpretation

- The effect of income is nearly always significant, though it gets smaller as education gets bigger. That is, as education increases, we expect smaller increases in prestige from increasing income
- The effect of education is significant and positive until around 16,000 dollars, which is around 2/3 the range of `income`, but is the 96^{th} percentile because of the skewness of income.
- This suggests that people tend to derive prestige from either higher incomes or higher education, but not really both.

Notes

Type notes here...

When Confidence Bounds Equal Zero

You may want to know when the confidence bounds are equal to zero. Consider the equation:

$$\hat{Y}_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i1} X_{i2}$$

- We know that the conditional effect of X_1 is $\beta_1 + \beta_4 X_2$ and that the lower bound is $(\beta_1 + \beta_4 X_2) - 1.96 \times SE(\beta_1 + \beta_4 X_2)$.
- Since those are all quantities that we know (or estimate), we could set it equal to zero and solve.
- This is what the `changeSig` function does.

Notes

Type notes here...

Change in Significance

```r
changeSig(mod, c("income", "education"))
```

```
## LB for B(income | education) = 0 when education=15.4979 (95th pctile)
## UB for B(income | education) = 0 when education=27.9396 (> Maximum Value in Data)
## LB for B(education | income) = 0 when income=15.9273 (96th pctile)
## UB for B(education | income) = 0 when income=59.5175 (> Maximum Value in Data)
```

Notes

Type notes here...

Alternate Visualization

An alternate way to visualize the information is with a three-dimensional surface.

```
DAintfun(mod, c("income","education"),
        theta=-45, phi=20)
```

Notes

Type notes here...

BGM Test for Prestige model

Here is the set of tests that Berry, Golder and Milton (2012) suggest. In the input to the function, the first variable in the `vars` argument is considered `X` and the second variable is considered `Z` for the purposes of the function.

```
##                est    se      t p-value
## P(X|Zmin)  2.445 0.520  4.698 0.000
## P(X|Zmax)  0.429 0.287  1.495 0.138
## P(Z|Xmin)  4.756 0.712  6.681 0.000
## P(Z|Xmax) -0.335 1.466 -0.229 0.820
## P(XZ)     -0.210 0.070 -3.012 0.003
```

Notes

Type notes here...

Centering and Interactions

- Let's assume we have the following model:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon_i$$

$$\beta = \begin{bmatrix} 2 \\ 3 \\ -4 \\ 3 \end{bmatrix}, X \sim \mathcal{N}_2(\mu, \boldsymbol{\Sigma}), \mu = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \Sigma = \begin{bmatrix} 1.0 & 0.4 \\ 0.4 & 1.0 \end{bmatrix}$$

- Both X variables are always positive and correlated at a reasonable level. Let's see what happens to the fitted values and coefficients when we mean-center them.

Notes

Type notes here...

Mean Centering

```
##
## ============================================================
##                                    Dependent variable:
##                                 ----------------------------
##                                              Y
##                                   Not Cent         Cent
##                                      (1)            (2)
## ------------------------------------------------------------
## x1                                 0.691         32.902***
##                                   (1.177)         (0.135)
##
## x2                               -6.074***       26.137***
##                                   (1.178)         (0.135)
##
## x1:x2                             3.221***        3.221***
##                                   (0.117)         (0.117)
##
## Constant                          23.554**       -1.287***
##                                   (11.746)        (0.132)
##
## ------------------------------------------------------------
## Observations                       1,000           1,000
## R2                                 0.994           0.994
## Adjusted R2                        0.994           0.994
## Residual Std. Error (df = 996)     3.910           3.910
## F Statistic (df = 3; 996)      53,680.490***  53,680.490***
## ============================================================
## Note:                          *p<0.1; **p<0.05; ***p<0.01
```

Notes

Type notes here...

VIF Statistics

	No Cent	Cent
x1	90.54	1.19
x2	90.73	1.19
x1:x2	251.45	1.00

Notes

Type notes here...

Conditional Effect of X

- Since we've moved the X's around, we need to consider not the effects in the model, but the conditional effects holding the X's at the same places *relative* to their respective distributions, for instance:

	x1	x1 (cent)	x2	x2 (cent)
25th	9.34	-0.66	9.31	-0.69
50th	10.00	-0.00	10.02	0.02
75th	10.64	0.64	10.71	0.71

Notes

Type notes here...

Conditional Effect of X(2)

- Now, we can look at the conditional effects of X_1 and X_2 at the given values above:

	eff x1	eff x1 (cent)	eff x2	eff x2 (cent)
25th	30.68	30.68	24.03	24.03
	0.16	0.16	0.15	0.15
50th	32.96	32.96	26.13	26.13
	0.14	0.14	0.14	0.14
75th	35.18	35.18	28.21	28.21
	0.16	0.16	0.15	0.15

Conditional Effects of x1 and x2

Notes

Type notes here...