Multiple Predictors in Linear Models

Dr. J. Kyle Roberts

Southern Methodist University
Simmons School of Education and Human Development
Department of Teaching and Learning

Adding a Second Predictor

- First reconsider our original model for a single predictor where:

$$y_i = a + b * x_i + \epsilon_i \quad i = 1, \ldots, n.$$

  and

$$\epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

- Were we to add another predictor to the above model, this would change the model to:

$$y_i = a + b_1 * x_1 + b_2 * x_2 + \epsilon_i.$$

  where b_1 represents the partial regression coefficient for y on x_1 when x_2 is in the equation and b_2 represents the partial regression coefficient for y on x_2 when x_1 is in the equation.

Predictive Coefficients and Multiple R^2

- In the case where we are using two covariates to predict a response on a single dependent variable, we can think of the product of these coefficients as:

$$\hat{y} = a + b_1 * x_1 + b_2 * x_2.$$

- $\hat{y}$ is considered the predicted score on y given the two covariates x_1 and x_2.
- The Multiple R^2 can then be thought of as:

$$R^2_{y.12} = \frac{r^2_{yx_1} + r^2_{yx_2} - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r^2_{x_1x_2}}$$

or

$$R^2_{y.12} = R^2_{y\hat{y}}$$

Transforming Between b and β

- We can easily go back and forth between b and β with

$$b = \beta(SD_y/SD_x)$$

or

$$\beta = b(SD_x/SD_y)$$

- In this case, b will equal β when the SDs are both 1.
- When the two variables are perfectly uncorrelated, then:

$$r_{xy} = \beta = b = 0$$

- Interpretation of β weights will be discussed further when we look at structure coefficients.

β Weights

- The weights from the original linear model represent the unstandardized weights (or weights in the metric of the original variables). In multiple regression, the β weights represent a standardized regression weight (or weights that are in z-score form).
- These may be thought of as regression weights when all variables are in z-score form or:

$$\hat{z}_y = \beta_{y1.2} z_1 + \beta_{y2.1} z_2$$

- The individual β weights may also be interpreted as

$$\beta_{y1.2} = \frac{r_{y1} - r_{y2} r_{12}}{1 - r_{12}^2}$$

$$\beta_{y2.1} = \frac{r_{y2} - r_{y1} r_{12}}{1 - r_{12}^2}$$

Computing β Weights

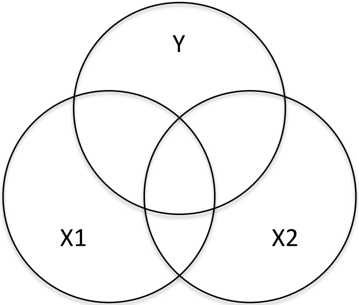
```
> data1 <- data.frame(x = 1:10, y = c(1:5, 8:12),
+      z = c(1:7, 11, 13, 19))
> (m1 <- lm(y ~ x + z, data1))

Call:
lm(formula = y ~ x + z, data = data1)

Coefficients:
(Intercept)            x            z
   -0.87558      1.44747     -0.08247

> data2 <- data.frame(x = scale(1:10), y = scale(c(1:5,
+      8:12)), z = scale(c(1:7, 11, 13, 19)))
> (m2 <- lm(y ~ x + z, data2))

Call:
lm(formula = y ~ x + z, data = data2)

Coefficients:
(Intercept)            x            z
 -6.912e-17    1.101e+00   -1.170e-01
```

Structure Coefficients

- Structure coefficients represent the correlation between a predictor variable and the predicted scores given the full regression equation.

$$r_s = r_{x \text{ with } y}/R = r_{x\hat{y}}$$

- There will be as many structure coefficients as there are dependent variables.



Computing Structure Coefficients

```
> m1

Call:
lm(formula = y ~ x + z, data = data1)

Coefficients:
(Intercept)            x            z
   -0.87558      1.44747     -0.08247

> cor(m1$fitted, data1$x)

[1] 0.9991804

> cor(m1$fitted, data1$z)

[1] 0.9245726
```

Explaining the Unique Contribution of Each Covariate

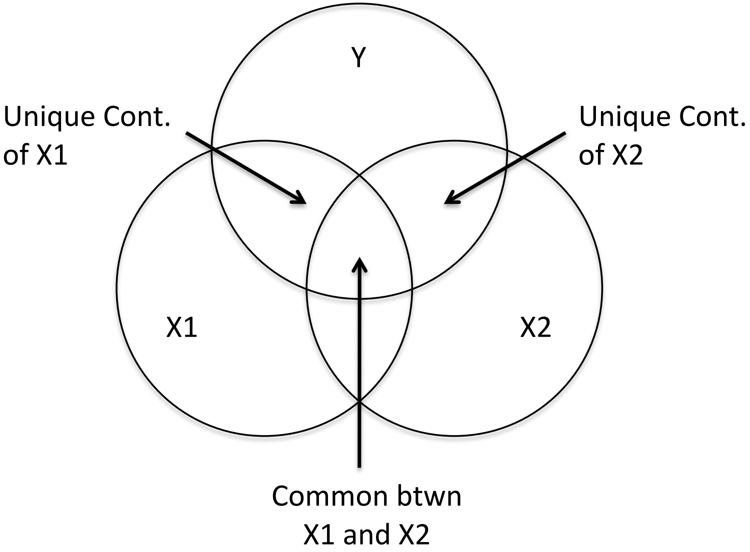
- From a given linear model with two predictors, the explained variance ($R^2_{0.12}$) can be partitioned into three components

$$
\begin{aligned}
\gamma_1 &= \text{unique contribution of } X_1 \text{ to } R^2_{0.12} \\
\gamma_2 &= \text{unique contribution of } X_2 \text{ to } R^2_{0.12} \\
\gamma_{12} &= \text{common contribution of } X_1 \text{ and } X_2 \text{ to } R^2_{0.12}.
\end{aligned}
$$

- Alternatively, the unique and common contributions of each of the predictor variables could be expressed as:

$$
\begin{aligned}
\gamma_1 &= R^2_{0.12} - R^2_{02} \\
\gamma_2 &= R^2_{0.12} - R^2_{01} \\
\gamma_{12} &= R^2_{01} + R^2_{02} - R^2_{0.12}.
\end{aligned}
$$

Graphical Example of Commonality Coefficients



An R Example

- We will be working through the example in the `yhat` library. Download and install and then look at the `help` by typing `?regr`.
- `regr` is a piece of code written by Nimon and Roberts that transforms the typical `lm` output to include β weight, structure coefficients, and commonality coefficients.

Homework

1. Create a dataset with two independent variables that are very highly correlated with each other, but only mildly correlated with the dependent variable. Run your analysis with `regr` and note what is happening with the β weight, structure coefficients, and commonality coefficients.
2. Create a dataset with two independent variables that have a near zero correlation with each other. Let the dependent variable be `1:10`. Run your analysis with `regr` and note what is happening with the β weight, structure coefficients, and commonality coefficients.