

Neuendorf

Dummy Coding and Effects Coding

For a nominal variable with c categories, you may create up to $c-1$ dummy or effects variables.
For all equations below, assume the all-standardized situation.

Imagine a four-group religion variable: All respondents are either self-declared Christian, Muslim, Jewish, or Buddhist. From this single nominal variable, three dummies (or effects coded variables) may be created.

DUMMY CODING

| | <u>D₁</u> | <u>D₂</u> | <u>D₃</u> | |
|-----------|----------------------|----------------------|----------------------|-------------------------------------|
| Christian | 1 | 0 | 0 | |
| Muslim | 0 | 1 | 0 | |
| Jewish | 0 | 0 | 1 | |
| Buddhist | 0 | 0 | 0 | – [The chosen “reference category”] |

$$\begin{aligned}
 \text{Christian } Y' &= \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 \\
 &= \beta_1 D_1 + 0 + 0 \\
 &= \beta_1 D_1 \\
 &= \beta_1 \quad \quad \quad \text{[because } D_1 = 1\text{]}
 \end{aligned}$$

$$\begin{aligned}
 \text{Muslim } Y' &= \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 \\
 &= 0 + \beta_2 D_2 + 0 \\
 &= \beta_2 D_2 \\
 &= \beta_2 \quad \quad \quad \text{[because } D_2 = 1\text{]}
 \end{aligned}$$

$$\begin{aligned}
 \text{Jewish } Y' &= \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 \\
 &= 0 + 0 + \beta_3 D_3 \\
 &= \beta_3 D_3 \\
 &= \beta_3 \quad \quad \quad \text{[because } D_3 = 1\text{]}
 \end{aligned}$$

$$\begin{aligned}
 \text{Buddhist } Y' &= \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 \\
 &= 0 + 0 + 0 \\
 &= 0
 \end{aligned}$$

| <u>Original Category</u> | <u>Expected value (Y')</u> |
|--------------------------|----------------------------|
| Christian | β_1 |
| Muslim | β_2 |
| Jewish | β_3 |
| Buddhist | 0 |

Hence:

- β_1 is the difference between the expected values of Christian and Buddhist,
- β_2 is the difference between the expected values of Muslim and Buddhist, and
- β_3 is the difference between the expected values of Jewish and Buddhist.

Compare the meaning of the partial regression coefficients with the simple r 's between Y and D_1 , D_2 , and D_3 .

EFFECTS CODING

| | <u>E₁</u> | <u>E₂</u> | <u>E₃</u> | |
|-----------|----------------------|----------------------|----------------------|-------------------------------------|
| Christian | 1 | 0 | 0 | |
| Muslim | 0 | 1 | 0 | |
| Jewish | 0 | 0 | 1 | |
| Buddhist | -1 | -1 | -1 | - [The chosen "reference category"] |

$$\begin{array}{l}
 \text{Christian } Y' = \\
 \text{Muslim } Y' = \\
 \text{Jewish } Y' =
 \end{array}
 \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \text{All three same as for dummy coding}$$

$$\begin{aligned}
 \text{Buddhist } Y' &= \beta_1 E_1 + \beta_2 E_2 + \beta_3 E_3 \\
 &= \beta_1(-1) + \beta_2(-1) + \beta_3(-1) \\
 &= 0 - \beta_1 - \beta_2 - \beta_3
 \end{aligned}$$

| <u>Original Category</u> | <u>Expected value (Y')</u> |
|--------------------------|-----------------------------------|
| Christian | β_1 |
| Muslim | β_2 |
| Jewish | β_3 |
| Buddhist | $0 - \beta_1 - \beta_2 - \beta_3$ |

Hence:

- β_1 is ½ the difference between the expected values of Christian and the other three groups (Muslim, Jewish, Buddhist); the sum of the coefficients for those three groups is $-\beta_1$
- β_2 is ½ the difference between the expected values of Muslim and the other three groups (Christian, Jewish, Buddhist); the sum of the coefficients for those three groups is $-\beta_2$
- β_3 is ½ the difference between the expected values of Jewish and the other three groups (Christian, Muslim, Buddhist); the sum of the coefficients for those three groups is $-\beta_3$

Again, compare the meaning of the partial regression coefficients with the simple r's between Y and E1, E2, and E3.

IN SUM, THEN:

1. For dummy coding, each test of a β indicates the difference between the dummy group and the "reference category."
2. For effects coding, each test of a β indicates the difference between the effects group and all other groups pooled.
3. Also, to test for the impact of a set of dummies or effects variables, include all as a block and look at the size and significance of the R^2 change.