Indicator Variables With Stata

Andy Grogan-Kaylor

24 Feb 2021 09:26:43

# Introduction

# Get Data

. use https://www.stata-press.com/data/r16/margex, clear  
(Artificial data for margins)

# Descriptive Statistics

. summarize y  
  
 Variable │ Obs Mean Std. Dev. Min Max  
─────────────┼─────────────────────────────────────────────────────────  
 y │ 3,000 69.73357 21.53986 0 146.3

. tabulate sex  
  
 sex │ Freq. Percent Cum.  
────────────┼───────────────────────────────────  
 male │ 1,498 49.93 49.93  
 female │ 1,502 50.07 100.00  
────────────┼───────────────────────────────────  
 Total │ 3,000 100.00

. tabulate group  
  
 group │ Freq. Percent Cum.  
────────────┼───────────────────────────────────  
 1 │ 1,199 39.97 39.97  
 2 │ 1,118 37.27 77.23  
 3 │ 683 22.77 100.00  
────────────┼───────────────────────────────────  
 Total │ 3,000 100.00

# Regressions

## “Usual” Regression With Indicator Variables

. regress y i.sex i.group  
  
 Source │ SS df MS Number of obs = 3,000  
─────────────┼────────────────────────────────── F(3, 2996) = 152.06  
 Model │ 183866.077 3 61288.6923 Prob > F = 0.0000  
 Residual │ 1207566.93 2,996 403.059723 R-squared = 0.1321  
─────────────┼────────────────────────────────── Adj R-squared = 0.1313  
 Total │ 1391433.01 2,999 463.965657 Root MSE = 20.076  
  
─────────────┬────────────────────────────────────────────────────────────────  
 y │ Coef. Std. Err. t P>|t| [95% Conf. Interval]  
─────────────┼────────────────────────────────────────────────────────────────  
 sex │  
 female │ 18.32202 .8930951 20.52 0.000 16.57088 20.07316  
 │  
 group │  
 2 │ 8.037615 .913769 8.80 0.000 6.245937 9.829293  
 3 │ 18.63922 1.159503 16.08 0.000 16.36572 20.91272  
 │  
 \_cons │ 53.32146 .9345465 57.06 0.000 51.48904 55.15388  
─────────────┴────────────────────────────────────────────────────────────────

. est store M1 // store estimates

## Regression With No Constant and No Reference Category For One Independent Variable

. regress y i.sex ibn.group, noconstant  
  
 Source │ SS df MS Number of obs = 3,000  
─────────────┼────────────────────────────────── F(4, 2996) = 9162.52  
 Model │ 14772177 4 3693044.26 Prob > F = 0.0000  
 Residual │ 1207566.93 2,996 403.059723 R-squared = 0.9244  
─────────────┼────────────────────────────────── Adj R-squared = 0.9243  
 Total │ 15979744 3,000 5326.58132 Root MSE = 20.076  
  
─────────────┬────────────────────────────────────────────────────────────────  
 y │ Coef. Std. Err. t P>|t| [95% Conf. Interval]  
─────────────┼────────────────────────────────────────────────────────────────  
 sex │  
 female │ 18.32202 .8930951 20.52 0.000 16.57088 20.07316  
 │  
 group │  
 1 │ 53.32146 .9345465 57.06 0.000 51.48904 55.15388  
 2 │ 61.35908 .7006367 87.58 0.000 59.9853 62.73285  
 3 │ 71.96068 .7730326 93.09 0.000 70.44495 73.47641  
─────────────┴────────────────────────────────────────────────────────────────

. est store M2 // store estimates

## Compare These Approaches

. est table M1 M2, star  
  
─────────────┬────────────────────────────────  
 Variable │ M1 M2   
─────────────┼────────────────────────────────  
 sex │  
 female │ 18.322021\*\*\* 18.322021\*\*\*   
 │  
 group │  
 1 │ (base) 53.321461\*\*\*   
 2 │ 8.0376149\*\*\* 61.359076\*\*\*   
 3 │ 18.639222\*\*\* 71.960683\*\*\*   
 │  
 \_cons │ 53.321461\*\*\*   
─────────────┴────────────────────────────────  
 legend: \* p<0.05; \*\* p<0.01; \*\*\* p<0.001

## Display Combinations of Results With margins

. margins sex#group  
  
Adjusted predictions Number of obs = 3,000  
Model VCE : OLS  
  
Expression : Linear prediction, predict()  
  
─────────────┬────────────────────────────────────────────────────────────────  
 │ Delta-method  
 │ Margin Std. Err. t P>|t| [95% Conf. Interval]  
─────────────┼────────────────────────────────────────────────────────────────  
 sex#group │  
 male#1 │ 53.32146 .9345465 57.06 0.000 51.48904 55.15388  
 male#2 │ 61.35908 .7006367 87.58 0.000 59.9853 62.73285  
 male#3 │ 71.96068 .7730326 93.09 0.000 70.44495 73.47641  
 female#1 │ 71.64348 .6015065 119.11 0.000 70.46407 72.82289  
 female#2 │ 79.6811 .8022261 99.32 0.000 78.10813 81.25407  
 female#3 │ 90.2827 1.114023 81.04 0.000 88.09838 92.46703  
─────────────┴────────────────────────────────────────────────────────────────

## The noconstant Option Does *Not* Work With Two Indicator Variables

. regress y ibn.sex ibn.group, noconstant  
note: 3.group omitted because of collinearity  
  
 Source │ SS df MS Number of obs = 3,000  
─────────────┼────────────────────────────────── F(4, 2996) = 9162.52  
 Model │ 14772177 4 3693044.26 Prob > F = 0.0000  
 Residual │ 1207566.93 2,996 403.059723 R-squared = 0.9244  
─────────────┼────────────────────────────────── Adj R-squared = 0.9243  
 Total │ 15979744 3,000 5326.58132 Root MSE = 20.076  
  
─────────────┬────────────────────────────────────────────────────────────────  
 y │ Coef. Std. Err. t P>|t| [95% Conf. Interval]  
─────────────┼────────────────────────────────────────────────────────────────  
 sex │  
 male │ 71.96068 .7730326 93.09 0.000 70.44495 73.47641  
 female │ 90.2827 1.114023 81.04 0.000 88.09838 92.46703  
 │  
 group │  
 1 │ -18.63922 1.159503 -16.08 0.000 -20.91272 -16.36572  
 2 │ -10.60161 1.01299 -10.47 0.000 -12.58783 -8.615381  
 3 │ 0 (omitted)  
─────────────┴────────────────────────────────────────────────────────────────

## Display Combinations of Results With margins

. margins sex#group  
  
Adjusted predictions Number of obs = 3,000  
Model VCE : OLS  
  
Expression : Linear prediction, predict()  
  
─────────────┬────────────────────────────────────────────────────────────────  
 │ Delta-method  
 │ Margin Std. Err. t P>|t| [95% Conf. Interval]  
─────────────┼────────────────────────────────────────────────────────────────  
 sex#group │  
 male#1 │ 53.32146 .9345465 57.06 0.000 51.48904 55.15388  
 male#2 │ 61.35908 .7006367 87.58 0.000 59.9853 62.73285  
 male#3 │ 71.96068 .7730326 93.09 0.000 70.44495 73.47641  
 female#1 │ 71.64348 .6015065 119.11 0.000 70.46407 72.82289  
 female#2 │ 79.6811 .8022261 99.32 0.000 78.10813 81.25407  
 female#3 │ 90.2827 1.114023 81.04 0.000 88.09838 92.46703  
─────────────┴────────────────────────────────────────────────────────────────