

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2009, The Johns Hopkins University and John McGready. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.



JOHNS HOPKINS
BLOOMBERG
SCHOOL *of* PUBLIC HEALTH

Lecture 5d: Practice Problem Solutions

John McGready
Johns Hopkins University

CE Charges

1. The following MLR relates total charges from the carotid endarterectomy procedures to the subject's race. To start, the race is coded as 0 for Caucasians, 1 for African Americans, and 2 for others in neither of the first two race categories. Here is the distribution of race in the sample of persons from 1995 in Maryland who had such a procedure.

```
. tab race
```

Race	Freq.	Percent	Cum.
-----+-----			
Caucas	2,488	92.87	92.87
AfroAm	167	6.23	99.10
Other	24	0.90	100.00
-----+-----			
Total	2,679	100.00	

```
. tab race, nolab
```

Race	Freq.	Percent	Cum.
-----+-----			
0	2,488	92.87	92.87
1	167	6.23	99.10
2	24	0.90	100.00
-----+-----			
Total	2,679	100.00	

CE Charges

1. Now, here is the MLR using the “xi” option in Stata. Caucasians are the reference group.

```
. xi: regress totchg i.race
i.race          _Irace_0-2          (naturally coded; _Irace_0 omitted)

-----+-----
Source |           SS           df           MS           Number of obs =      2679
-----+-----
Model |  1.3918e+09           2       695906243       F(  2,  2676) =     25.91
Residual |  7.1861e+10       2676   26853776.6       Prob > F      =     0.0000
-----+-----
Total |  7.3253e+10       2678   27353442.4       R-squared     =     0.0190
                                           Adj R-squared =     0.0183
                                           Root MSE    =     5182.1

-----+-----
totchg |           Coef.      Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
_Irace_1 |    2934.388      414.2395     7.08  0.000     2122.127     3746.65
_Irace_2 |    1549.089     1062.874     1.46  0.145    -535.0484    3633.226
   _cons |    6888.995     103.8909    66.31  0.000     6685.28     7092.709
-----+-----
```

CE Charges

- a) What does the overall F-test conclude about a population-level association between total CE charges and race (assume a significance level of 0.05)?

```
. xi: regress totchg i.race
i.race          _Irace_0-2          (naturally coded; _Irace_0 omitted)

-----+-----
Source |           SS          df           MS              Number of obs =      2679
-----+-----
Model | 1.3918e+09           2       695906243          F( 2, 2676) =     25.91
Residual | 7.1861e+10        2676    26853776.6          Prob > F      =    0.0000
-----+-----
Total | 7.3253e+10        2678    27353442.4          R-squared     =    0.0190
                                           Adj R-squared =    0.0183
                                           Root MSE     =    5182.1

-----+-----
totchg |           Coef.      Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
_Irace_1 |    2934.388      414.2395       7.08  0.000     2122.127     3746.65
_Irace_2 |    1549.089     1062.874       1.46  0.145    -535.0484    3633.226
   _cons |    6888.995      103.8909     66.31  0.000     6685.28     7092.709
-----+-----
```

- The p-value from the F-test is very small ($< .001$): this leads to conclusions that at least one of the race group mean charges is statistically significantly different than at least one other race group.

CE Charges

- b) What is the estimated mean difference in CE charges for African-Americans compared to Caucasians?

```
. xi: regress totchg i.race
i.race      _Irace_0-2      (naturally coded; _Irace_0 omitted)
```

Source	SS	df	MS	Number of obs	=	2679
Model	1.3918e+09	2	695906243	F(2, 2676)	=	25.91
Residual	7.1861e+10	2676	26853776.6	Prob > F	=	0.0000
-----				R-squared	=	0.0190
-----				Adj R-squared	=	0.0183
Total	7.3253e+10	2678	27353442.4	Root MSE	=	5182.1

totchg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
_Irace_1	2934.388	414.2395	7.08	0.000	2122.127 3746.65
_Irace_2	1549.089	1062.874	1.46	0.145	-535.0484 3633.226
_cons	6888.995	103.8909	66.31	0.000	6685.28 7092.709

- With race coded 0 for Caucasians, this is the reference category as it is the lowest value of 0, 1, and 2. The coefficient for `_Irace_1` compares the mean charges for African Americans compared to the reference category.

CE Charges

c) What is the estimated mean CE charges for African-Americans?

```
. xi: regress totchg i.race
i.race          _Irace_0-2          (naturally coded; _Irace_0 omitted)

-----+-----
Source |           SS          df           MS              Number of obs =      2679
-----+-----
Model |  1.3918e+09           2       695906243          F( 2, 2676) =      25.91
Residual |  7.1861e+10        2676    26853776.6          Prob > F      =      0.0000
-----+-----
Total |  7.3253e+10        2678    27353442.4          R-squared     =      0.0190
                                           Adj R-squared =      0.0183
                                           Root MSE     =      5182.1

-----+-----
totchg |           Coef.      Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
_Irace_1 |  2934.388          414.2395       7.08  0.000     2122.127     3746.65
_Irace_2 |  1549.089          1062.874       1.46  0.145    -535.0484    3633.226
_cons |  6888.995          103.8909      66.31  0.000     6685.28     7092.709
-----+-----
```

— The estimated mean CE charges for African Americans is as follows: \$6,889 + \$2,934 = \$9,823

CE Charges

- d) What is the estimated mean difference in CE charges for African-Americans compared to non-Caucasians/non-African Americans?

```
. xi: regress totchg i.race
i.race          _Irace_0-2          (naturally coded; _Irace_0 omitted)

-----+-----
Source |           SS           df           MS           Number of obs =      2679
-----+-----+-----+-----+-----
Model |  1.3918e+09           2           695906243           F( 2, 2676) =      25.91
Residual |  7.1861e+10          2676           26853776.6           Prob > F      =      0.0000
-----+-----+-----+-----+-----
Total |  7.3253e+10          2678           27353442.4           R-squared     =      0.0190
                                           Adj R-squared =      0.0183
                                           Root MSE    =      5182.1

-----+-----
totchg |           Coef.      Std. Err.      t      P>|t|      [95% Conf. Interval]
-----+-----+-----+-----+-----+-----
_Irace_1 |    2934.388         414.2395       7.08   0.000       2122.127       3746.65
_Irace_2 |    1549.089        1062.874       1.46   0.145       -535.0484      3633.226
   _cons |    6888.995         103.8909      66.31   0.000       6685.28       7092.709
-----+-----
```

- This can be estimated by taking the slope for African Americans (_Irace_1) and subtracting the slope for other races (_Irace_2):
 $\$2,934 - \$1,549 = \$1,385$