JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Section 4e: Practice Problem Solutions

John McGready
Johns Hopkins University

Using Stata: Arm Circumference and Height

1. Recall the regression relating arm circumference to height for the random sample of 150 Nepali children less than 12 months old

```
. regress armcirc height

      Source |       SS       df       MS              Number of obs =     150
-------------+------------------------------           F(  1,   148) =  124.30
       Model |  148.874597     1  148.874597            Prob > F      =  0.0000
    Residual |  177.263335   148  1.19772523           R-squared     =  0.4565
-------------+------------------------------           Adj R-squared =  0.4528
       Total |  326.137932   149  2.18884518           Root MSE      =  1.0944

------------------------------------------------------------------------------
     armcirc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      height |   .1579469   .0141671    11.15   0.000     .1299511    .1859428
       _cons |   2.695906   .8774225     3.07   0.003     .9620116      4.4298
------------------------------------------------------------------------------
```

   - Suppose arm circumference had been reported in inches instead of centimeters, but height was still recorded in centimeters

Example: Arm Circumference and Sex

a) Can you determine what the resulting slope estimate for height would be with this different reporting schema?

   - *In the results given, both arm circumference and height are recorded in cm. The units of the slope are "cm AC per cm height." If arm circumference is instead recorded in inches and the regression results re-computed with this change, the resulting slope will be in units of "inches AC per cm height." How to convert the current slope of 0.16 cm AC/cm height to the new units? Well one inch = 2.54 cm; so the updated slope estimated will be 0.16 cm/ 2.54 cm/inch: 0.06 inches AC per cm height.*

Example: Arm Circumference and Sex

b) What would the R^2 value be with this different reporting schema?

- *R^2 would be unaffected by the change in units and would still be 0.46*

Example: Arm Circumference and Sex

2. Recall the regression relating arm circumference to child's sex for the random sample of 150 Nepali children less than 12 months old

$$\hat{y} = 12.5 + -0.13x$$

- R^2 for this analysis was 0.002 or 0.2%
  - In this example, x is the binary variable for sex, coded as a 1 for female children, and 0 for male children. Suppose x was coded as 1 for male children and 0 for female children.

a) What would the estimate of R^2 be?

R^2 would remain the same at 0.002 or 0.2%.

b) What would the estimate of *r* be?

Since the slope will be positive with the change in coding,

$$r = +\sqrt{R^2} = +\sqrt{0.002} \approx 0.045.$$