William Kuhns
4/22/2019
Aaron Meininger

1. Verbatim from the the book preface, “
MLR.1: Introduce the population model and interpret the population parameters (which we hope to estimate).
MLR.2: Introduce random sampling from the population and describe the data that we use to estimate the population parameters.
MLR.3: Add the assumption on the explanatory variables that allows us to compute the estimates from our sample; this is the so-called no perfect collinearity assumption.
MLR.4: Assume that, in the population, the mean of the unobservable error does not depend on the values of the explanatory variables; this is the “mean independence” assumption combined with a zero population mean for the error, and it is the key assumption that delivers unbiasedness of OLS.
MLR.5 (homoskedasticity) is added for the Gauss-Markov Theorem (and for the usual OLS variance formulas to be valid), (Wooldridge, xiv).“

   My brief interpretations of each assumption:
   SLR1 : We are making a linear unbiased relationship, setting up the equation, and defining our variables.
   SLR2 : The sample is random and of size n.
   SLR3 : The data varies. X the independent variable can’t be the same throughout the data set.
   SLR4: The mean error is 0 across the entire sample.
   SLR5: Doesn't matter what the input is the variance of the error should be the same.

2.

   The null hypothesis could be a bad guess. The sample mean could be subject to selection or non-response bias.
   Steps:
   - Make an initial Null Hypothesis
   - Get data’s mean and standard error
   - Get test statistic which is the difference between the sample mean and your null hypothesis over the standard error
   - Choose a significance level and find the degrees of freedom which is one minus your sample size
   - Refer to the t-table to see if the t-stat falls within the significance level, if it does fail to reject, and if it is outside the range then reject the hypothesis

3.

   A. The standard deviation of the mean annual income is 3045.

B. t-Stat = 1.7146, this falls outside our significance level of 90% or 1.292 at 99 df. Rejecting this hypothesis.

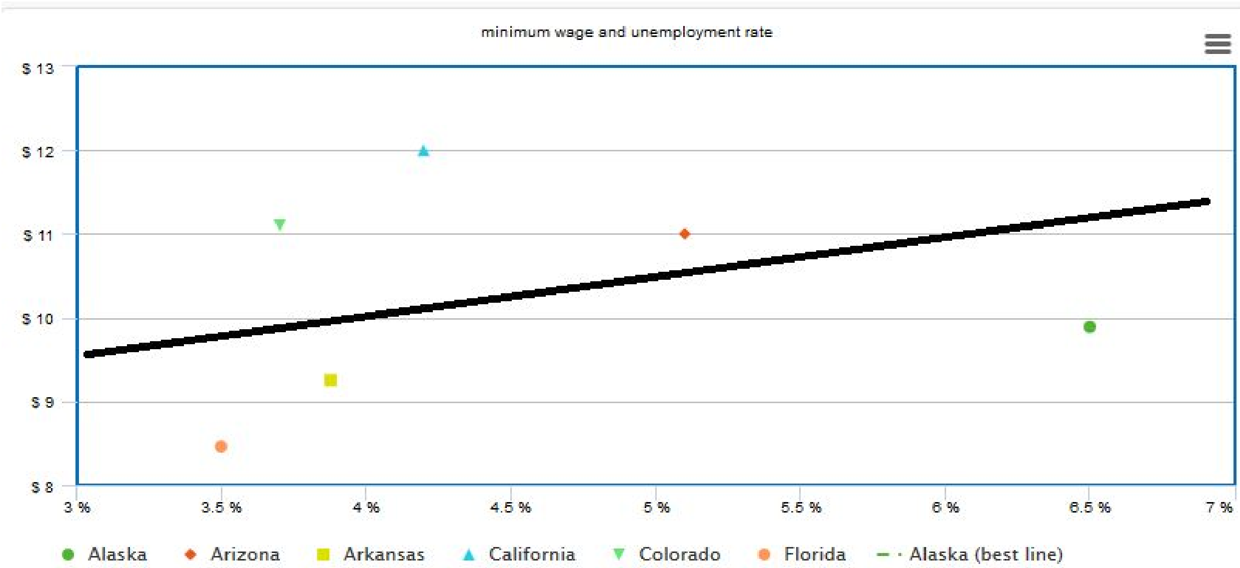
C. t-Stat = 1.0578, this fall inside our significance level of 95% or 1.990 at 99 df. Failing to Reject this hypothesis.

D. 39161.45 - 51280.55 at 95% confidence

E. P-value = 0.0432, using calculator

F. I failed to reject the null hypothesis of $42,000 due to it falling within the confidence interval.

4. A.



Y-intercept: $9.7 Slope: 1/2

B. Var(unemployment) = 1.10, Var(minimum wage) = 1.45,
Cov(unemployment, minimum wage) = 0.204

C. unemployment = x, minimum wage = y
X Mean = 4.467
Y Mean = 10.283

B_1 = cov(x,y)/var(x) = .204/1.1 =0.19

B_0 = 10.283 - 4.467(0.19) = 9.43

Y = 0.19x + 9.43, 0.19 is the slope

D. Y = 9, x= -2.26

E. unemployment rate will decrease by about 3/5.

|

$$R^2 = \frac{\sum (Y_i' - \overline{Y})^2}{\sum (Y_i - \overline{Y})^2}$$

F.
R^2 = 0.76

G. The model needs to be linear. No, regression through the origin does not make any sense for this data. It would break linearity due to living standards. This would occur way before the minimum wage is zero.

H.

$$\sum_{i=1}^{n}(y_i - \tilde{\beta}_1 x_i)^2. \qquad \tilde{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2},$$

B_1 = 2.19
y=2.19x

I. y=9, x =4.11 This answer makes sense the answer in D does not. You can not have a negative unemployment rate.

5. A. that is the slope, for every new park the house prices increases by 9000
B. 870000, 1050000
C. 76 parks
D. 259000 off or 26.73% error
E. 15,900,000, this number is too large because New York City is fairly large, the number of parks should only include local parks and the condition of the parks.
F. That’s the slope, this now factors in the density of the population in the area.
G. 30/20 * 53100 + 510000 = 589,650
H. 17/80 * 53100 + 510000 = 521,283.75
I. e^(0.7 * parksper1000HHs) is now the new independent variable. The 0.7 is used to scale the effect parks have on house prices. It’s still a slope just not a linear one.
J. Not much, e^(0.7) = 2.01, e^(1.05) = 2.85

6.
A.
regress colGPA hsGPA

Source | SS df MS Number of obs = 141

```
-------------+----------------------------------   F(1, 139)      =      28.85
       Model |  3.33506006          1  3.33506006   Prob > F      =     0.0000
    Residual |  16.0710394        139  .115618988   R-squared     =     0.1719
-------------+----------------------------------   Adj R-squared  =    0.1659
       Total |  19.4060994    140  .138614996   Root MSE      =     .34003

------------------------------------------------------------------------------
      colGPA |      Coef.   Std. Err.      t      P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       hsGPA |   .4824346   .0898258     5.37   0.000     .304833    .6600362
       _cons |   1.415434   .3069376     4.61   0.000     .8085635   2.022304
------------------------------------------------------------------------------
```

B. 1.415434 is the intercept, and .482 is the slope, this makes sense as HS gpa does have some impact on college gpa but not a lot.

C. about half at .75

D. 50%

7.

```
A.   Model |  1.33028272
B.  Y = -.00298x + 3.15
C.  X = 20, y = 3.01
D.
E.          Source |        SS             df       MS     Number of obs   =     141
F.  -------------+----------------------------------   F(1, 139)      =      10.23
G.          Model |  1.33028272          1  1.33028272   Prob > F      =     0.0017
H.          Residual |  18.0758167        139  .130041847   R-squared     =     0.0685
I.  -------------+----------------------------------   Adj R-squared  =    0.0618
J.          Total |  19.4060994    140  .138614996   Root MSE      =     .36061
K.
L.  ------------------------------------------------------------------------------
M.          colGPA |      Coef.   Std. Err.      t      P>|t|     [95% Conf. Interval]
N.  -------------+----------------------------------------------------------------
O.  lecturessk~d |   -.002984      .000933     -3.20   0.002   -.0048287   -.0011394
P.          _cons |   3.153084   .0427751     73.71   0.000   3.06851      3.237658
Q.  ------------------------------------------------------------------------------
```

y= -.01x + 3.15

The slope is steep in part D.

.