CMU MSP 36602 Review of Regression Interpretation HJS, Jan 14, 2019

1) If the standard regression assumptions hold, then:

a. $E(Y_i | \mathbf{x_i}, \boldsymbol{\beta}) =$

b. $V(Y_i | \mathbf{x_i}, \boldsymbol{\beta}, \sigma^2) =$

c. $E(\widehat{\boldsymbol{\beta}} | \mathbf{X}) =$

d. $V(\widehat{\boldsymbol{\beta}} | \mathbf{X}, \sigma^2) =$

e. $\widehat{Y_i} | \boldsymbol{x_i}, \widehat{\boldsymbol{\beta}} =$

f. $E(\widehat{Y_i} | \boldsymbol{x_i}, \widehat{\boldsymbol{\beta}}) =$

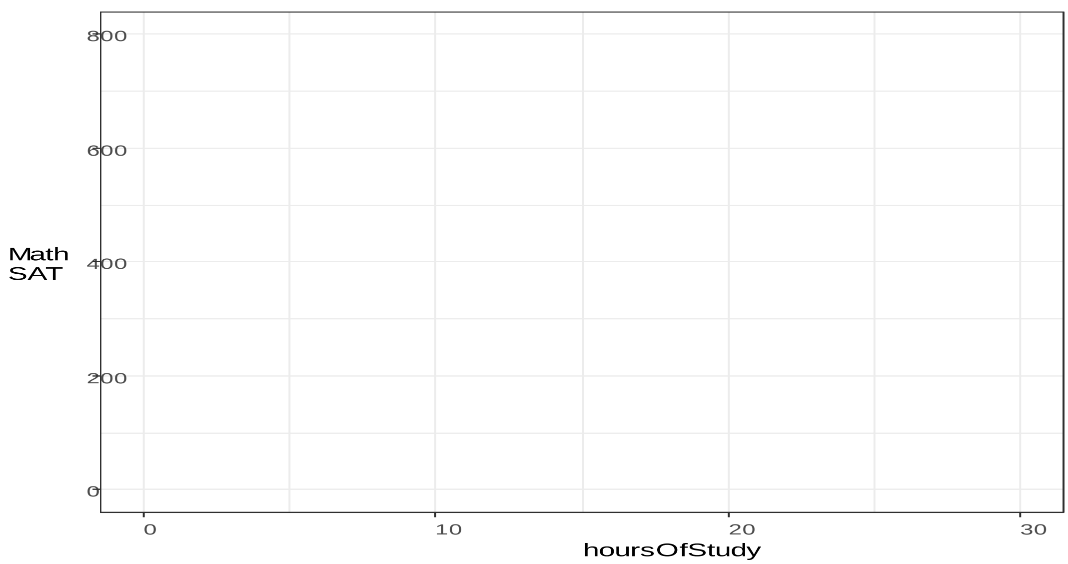
g. $V(\widehat{Y_i} | \boldsymbol{x_i}, \widehat{\boldsymbol{\beta}}) =$

2) Picturing the means model

a. Plot the model for an SAT Math improvement (in points out of 800) corresponding to:

Source	Estimate	SE	p-value
(Intercept)	640	5.0	<0.0001
MethodB	-12.0	2.0	<0.0001
HoursOfStudy	3.0	1.0	0.007

where HoursOfStudy ranges from 0 to 20 and the other method is "A".

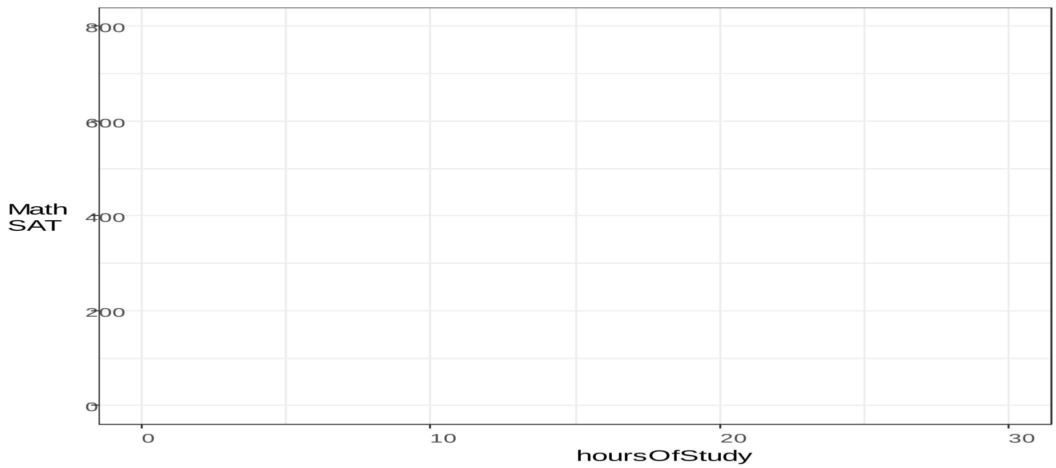


Interpretations:

What about if $\hat{\beta}_{\text{method}}$ = -2.0 with SE=22.0 and p=0.93?

b. Plot the model for an SAT Math improvement (in points out of 800) corresponding to:

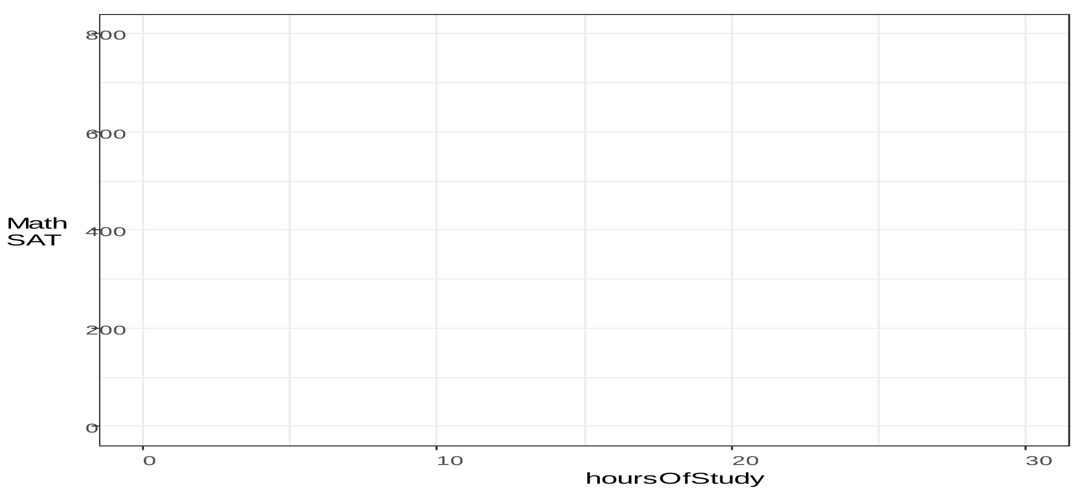
Source	Estimate	SE	p-value
(Intercept)	640	5.0	<0.0001
MethodB	12.0	2.0	<0.0001
HoursOfStudy	3.0	1.0	0.007
MethodB:HoursOfStudy	-6.0	2.0	0.007



Interpretations:

c. Plot the model for an SAT Math improvement (in points out of 800) corresponding to:

Source	Estimate	SE	p-value
(Intercept)	520	5.0	<0.0001
MethodB	-2.0	0.4	<0.0001
HoursOfStudy	32.0	3.0	<0.0001
HoursOfStudySquared	-1.1	0.3	0.002



Interpretations:

3) Adding uncertainty

a) Using model *2a)*, write out the equation for standard error of the estimate of the mean of y when **x**=(1, x_1, x_2).

b) What would it look like on the plot?

c) What do we need to do to get prediction interval rather than confidence intervals for the mean? What assumption(s) become more important?