Self-Assessment
Weeks 2 and 3: Regression with One Quantitative Predictor; Model Fit

Regression

1. Below is a regression analysis with two variables:

DV = violent crime rate per 100,000 in each US state
IV = percent of US state population living in single-parent households

Descriptive statistics for the two variables are provided first. Violent crime rate ranges from a low of 82 (North Dakota) to a high of 1206 (Florida), with a mean crime rate of 566.67. Percent of population in single-parent households ranges from a low of 8.4% (North Dakota) to a high of 14.9% (Louisiana), with a mean of 11.11%.

Statistics

		violent crime rate per 100,000	percent of population in single parent family
N	Valid	50	50
	Missing	0	0
Mean		566.66	11.1100
Median		509.50	10.9000
Mode		208	10.80
Std. Deviation		295.877	1.47513
Variance		87543.372	2.176
Range		1124	6.50
Minimum		82	8.40
Maximum		1206	14.90

Regression results are presented below.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.649(a)	.421	.409	227.514

a Predictors: (Constant), percent of population in single parent family

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1805011.036	1	1805011.036	34.871	.000(a)
	Residual	2484614.184	48	51762.796		
	Total	4289625.220	49			

a Predictors: (Constant), percent of population in single parent family
b Dependent Variable: violent crime rate per 100,000

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-878.861	246.895		-3.560	.001	-1375.278	-382.445
	percent of population in single parent family	130.110	22.033	.649	5.905	.000	85.809	174.411

a Dependent Variable: violent crime rate per 100,000

The prediction equation for this regression model is

Predicted Y = b0 + b1 X

Predicted Crime Rate per 100,000 = -878.86 + 130.11 (Percent in Single-parent Households)

(a) What is the literal interpretation for values obtained for b0 and b1?
(b) What is the general interpretation for b1?
(c) What percent of the crime rate variance can be predicted by knowing the percent of households that are single parents?
(d) Does this model predict more variance in crime rates than would be expected by chance?
    (d1) Which is the statistic used to measure how much variance is predicted, and what amount was predicted for this model?
    (d2) What is the symbolic and written null hypothesis assessed in question (d)?
    (d3) What is the test statistic used to test the null hypothesis found in question (d2)?
    (d4) What p-value is reported in the regression analysis for the test statistic sought in question (d3)?
(e) Is the slope, b1, for this model statistically significant at the .05 level (explain your response)?
(f) What is the interpretation for the 95% confidence interval for b1?
(g) How can the 95% confidence interval for b1 be used to test Ho: b1 = 0.00?
(h) What is the predicted violent crime rate for North Dakota (8.4% of households are single parents)?
(i) What is the predicted violent crime rate for Louisiana (14.9% of households are single parents)?
(j) What would be the residual for North Dakota (observed violent crime rate for ND = 82)?
(k) What would be the residual for Louisiana (observed violent crime rate for LA = 1062)?

2. Using the blood pressure data from the Week 1 self-assessment, perform a regression analysis in which the IV = body weight (pounds) and DV = systolic blood pressure.

SPSS Data: http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week01/Week01Q5Data.sav
Excel Data: http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week01/Week01Q5Data.xlsx

Present APA styled results for this analysis.

3. What does mean squared error (MSE, or mean squared residual, MSR) measure?

4. What does standard error of estimate (SEE, or standard error of residual, SER) measure?

5. If the mean squared error (MSE, or mean squared residual, MSR) is 25, what would be the standard error of estimate (SEE, or standard error of residual, SER)?

6. If the original variance of a dependent variable (DV) is 50, and the MSE is 35, what is the value of adjusted R^2?

7. There is a commonly used data file in statistics that contains a number of automobile measurements including miles per gallon (MPG) and vehicle weight in pounds. The data file contains over 400 records, but for this problem I selected 10 observations. I ran a regression analysis with

IV = vehicle weight in pounds
DV = miles per gallon (MPG)

Below is the SPSS coefficient table for this analysis.

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	52.519	6.836		7.683	.000	36.756	68.282
	Vehicle Weight (lbs.)	-.011	.003	-.811	-3.919	.004	-.017	-.004

a Dependent Variable: Miles per Gallon

The prediction equation is as follows:

Predicted Y = b0 + b1 X

Predicted MPG = 52.52 + (-0.011 x vehicle weight)

Below is at table that shows the observed MPG and the predicted MPG. The predicted MPG values were obtained using the equation above. The values for vehicle weight have been removed from the table.

Predicted MPG Table

mpg	weight	predicted mpg
15	---	11.897
17	---	14.581
19	---	23.546
22	---	26.032
25	---	29.134
28	---	27.352
29	---	31.983
32	---	32.324
36	---	32.445
43	---	30.685

Using the information provided in the Predicted MPG Table, calculate and report two model fit statistics, R and R^2.