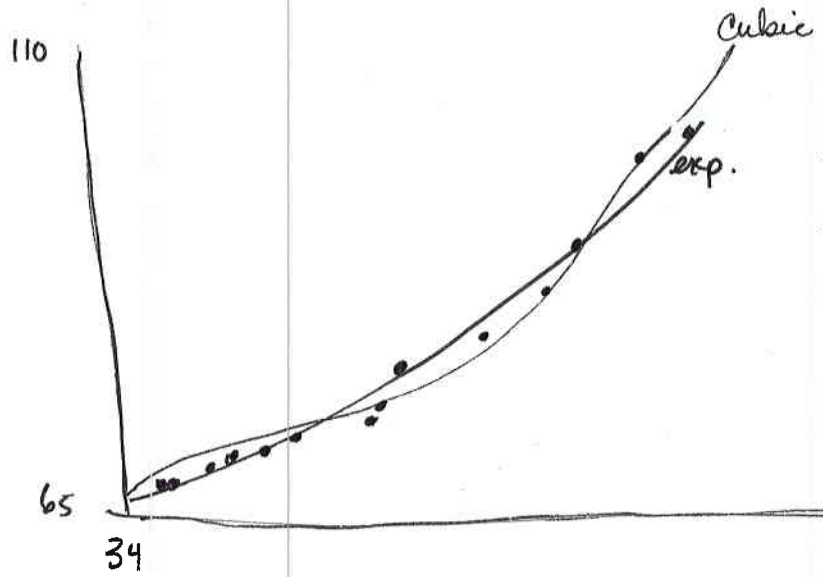
Stat 2470, Quiz #14, Spring 2015 Name KEY

Instructions: Show all work. State any formulas used. If you use the calculator, you should say which function you used, and what you entered into it, as well as any output. I can only give partial correct for incorrect answers if I have something to grade.

1. Use the data to find an exponential regression model. Plot the data on a scatterplot along with the regression equation. State the correlation for the model.

x	39	40	45	47	50	53	62	63	65	73	73	78	80	82	89
y	71	71	73	74	75	76	78	80	84	85	85	88	93	101	104



$y = a \cdot b^x$

$y = 52.3228(1.007)^x$

$r = .9609$

$r^2 = .923$

2. Calculate a cubic regression equation for this model. State the R^2 value. Is it better or worse than the exponential model? Which would you choose to model the data? [Hint: Remember to take into account that the fewer parameters that explain pretty well, the more likely the model is to hold up with the collection of additional data.]

$y = 2.71x^3 + (-.0397)x^2 + 2.273x + 26.527$

$R^2 = .96247$

I would choose the exponential model. It has fewer parameters and a pretty good fit. The cubic one has twice the parameters.

3. Let y be sales of a fast-food outlet (in thousands of dollars), with x_1 the number of competing outlets within a 1-mile radius, x_2 the population (in thousands) within a 1-mile radius and x_3 and indicator variable to be 1 if there is a drive-thru window and 0 if there is not. Suppose the regression model is given as $y = 10 - 1.2x_1 + 6.8x_2 + 15.3x_3 + \epsilon$.

   a. What is the mean value of sales when the number of competing outlets is 2 and there are 10,000 people in the area, and the outlet does have a drive-thru window?

$10 - 1.2(2) + 6.8(10) + 15.3(1) = 90.9$

$90,900

   b. Interpret the coefficient on x_2 in the context of the problem. What does it mean for the value of y when other variables are held constant?

for each 1000 increase in population Sales goup $6,800 if all other things remain equal.

   c. If you were a potential franchise owner looking to purchase an existing store, one with 3 competing outlets in the area, 7000 people and a drive-thru window, versus one with 1 competing outlet, a population of 8500 people nearby, but no drive-thru window, use the regression equation to determine which outlet was likely to be a better value in terms of mean sales.

Outlet $y_1 = 10 - 1.2(3) + 6.8(7) + 15.3(1) = 69.3$

outlet $y_2 = 10 - 1.2(1) + 6.8(8.5) + 15.3(0) = 66.6$

I would go with outlet 1 as it has higher mean predicted sales despite the lower population and higher competition.