Economics 91 Spring 2010
Linus Yamane Pitzer College

Regression

1. Suppose a real estate broker wishes to build a predictive model evaluating the price of coastline housing tracts in southern California. From county records, she obtains a random sample of 15 properties and their selling price per lot (in hundreds of thousands of dollars). After visiting the properties, the broker assigns points on a scale of 0-5 for the property's views of the ocean. The higher the score, X_1, the more attractive the view. Additionally, she rates the property's distance from the ocean, X_2, in terms of city blocks. A property directly on the ocean is given a 0 value. The regression output is presented below. Standard errors are in parentheses.

$$\begin{array}{c} \text{Price} = 2.60 + 0.800\ \text{View} - 0.600\ \text{Dist} \\ (0.7521)\ \ (0.2047)\ \ \ (0.2534) \end{array}$$

a. Interpret the coefficient estimates. Do they have the expected sign?
b. A particular lot has a 3 rating for view, and the property is 2 blocks from the ocean. What is the estimated cost of this lot?
c. Test the hypothesis that distance from the beach does not matter for housing prices at the 5% level.
d. If the standard deviation of View is 1.49 and the standard deviation of Dist is 1.89, which variable has more of an influence on price? Explain.

yes,3.8,reject,view

2. An aircraft company wanted to predict the number of worker-hours necessary to finish the design of a new plane. Relevant explanatory variables were thought to be the plane's top speed, its weight, and the number of parts it had in common with other models built by the company. A sample of 27 of the company's planes were taken, and the follow regression estimated:

$$\begin{array}{c} Y = 7.35 + 0.661\ SP + 0.065\ WT - 0.018\ CP \\ (1.23)\ \ (0.099)\ \ \ \ (0.032)\ \ \ \ (0.0023) \end{array}$$

where
Y = design effort, in millions of worker-hours
SP = plane's top speed, in Mach number
WT = plane's weight, in tons
CP = percentage number of parts in common with other models ($0 \le CP \le 1$)
and estimated standard errors appear in parentheses.

a. Interpret the coefficient estimates. Are the coefficient signs as you would expect?
b. How many worker hours would you expect to need to design a plane with a top speed of Mach 5 which weighs 4 tons and has 30% of its parts in common with other models?
c. Test the null hypothesis that, all else being the same, the plane's weight has no linear influence on design effort against the two-sided alternative at the 5% level.

d. Test the null hypothesis that, all else being the same, an increase in the plane's top speed by 1 Mach number increases the design effort by 0.50 million worker-hours against the one-sided alternative that it increases design effort by more than 0.50 million worker-hours at the 1% level.
yes, 10.9096,accept,accept

3. A random sample of 17 golfers on the PGA tour at the end of 1993 recorded their total earnings for the year (in multiples of $100,000), their average driving distance (in yards), and the average number of putts per hole. The multiple linear regression model was estimated as follows:

$$Earnings = \beta_0 + \beta_1(DrivingDist) + \beta_2(\#Putts) + \varepsilon_i$$

$$Earnings = 75.0182 + 0.3057(DrivingDist) - 84.2894(\#Putts)$$

$$(59.721)\ \ (0.086) \qquad\qquad (30.930)$$

a. Find the 90% confidence interval for β_1, the coefficient of driving distance.

b. You are interested in determining if putting performance has any effect on earnings. Test the hypothesis $H_0 : \beta_2 = 0$ versus the alternative hypothesis $H_a : \beta_2 \neq 0$. What do you conclude about putting performance?