

Homework - 4
CIVL 7012/8012: Probabilistic Methods for Engineers
Due: April 4, 2018

Question-1: (Dataset: TRT.txt). You are given 151 observations of a travel survey collected in State College, Pennsylvania. All of the households in the sample are making the morning commute to work. They are all departing from the same origin (a large residential complex in the suburbs) and going to work in the Central Business District. They have the choice of three alternate routes; 1) a four-lane arterial (speed limit = 35mph, 2 lanes each direction), 2) a two-lane rural road (speed limit = 35mph, 1 lane each direction) and 3) a limited access four-lane freeway (speed limit = 55mph, 2 lanes each direction). Your task is to estimate a model of individual average travel speed to work using standard regression techniques. Your solution to this problem should include:

- (i) The results of your best model specification.
- (ii) A discussion of the logical process that led you to the selection of your final specification. (e.g. Discuss the theory behind the inclusion of your selected variables). Include t-statistics and justify the sign of your variables.

Download data for question 1 [HERE](#).

Variables available for your specification are: (file *trt.out*)

Variable	Explanation
x1	Actual in-vehicle travel time in minutes
x2	Route chosen: 1 - arterial, 2 - rural road, 3 - freeway
x3	Traffic flow rate at time of departure in vehicles per hour
x4	Number of traffic signals on the selected route
x5	Distance along the selected route in tenths of miles
x6	Seat belts: 1 - if wear, 0 - if not
x7	Number of passengers in car
x8	Driver age in years: 1 - 18 to 23, 2 - 24 to 29, 3 - 30 to 39, 4 - 40 to 49, 5 - 50 and above
x9	Gender: 1 - male, 0 - female
x10	Marital status: 1 - single, 0 - married
x11	Number of children
x12	Annual income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - more than 50000
x13	Model year of car (e.g. 86 = 1986)
x14	Origin of car: 1 - domestic, 0 - foreign

Question-2: (Dataset: VOTE1.RAW). The following model can be used to study whether campaign expenditures affect election outcomes:

$$voteA = \beta_0 + \beta_1 \log(expendA) + \beta_2 \log(expendB) + \beta_3 prtystrA + u$$

where

voteA: is the percentage of the vote received by Candidate A,

expendA and *expendB*: are campaign expenditures by Candidates A and B,

prtystrA: is a measure of party strength for Candidate A (the percentage of the most recent presidential vote that went to A's party).

Determine the following.

- (i) What is the interpretation of β_1 ?
- (ii) In terms of the parameters, state the null hypothesis that a 1% increase in A's expenditures is offset by a 1% increase in B's expenditures.
- (iii) Estimate the given model using the data and report the results in usual form. Do A's expenditures affect the outcome? What about B's expenditures? Can you use these results to test the hypothesis in part (ii)?
- (iv) Estimate a model that directly gives the t statistic for testing the hypothesis in part (ii)?
- (v) What do you conclude? (Use a two-sided alternative.)

Download data for question 2 [HERE](#), and data description [HERE](#)

Question-3: Use the data in HPRICE1.RAW to estimate the model

$$price = \beta_0 + \beta_1 sqft + \beta_2 bdrms + u$$

where

price: is the house price measured in thousands of dollars,

sqft: house area in square feet

bdrms: number of bedrooms

- (i) Write out the results in equation form.
- (ii) What is the estimated increase in price for a house with one more bedroom, holding square footage constant?
- (iii) What is the estimated increase in price for a house with an additional bedroom that is 140 square feet in size? Compare this to your answer in part (ii).
- (iv) What percentage of the variation in price is explained by square footage and number of bedrooms?
- (v) The first house in the sample has *sqft* = 2,438 and *bdrms* = 4. Find the predicted selling price for this house from the OLS regression line.

- (vi) The actual selling price of the first house in the sample was \$300,000 (so price = 300). Find the residual for this house. Does it suggest that the buyer underpaid or overpaid for the house?

Download data for question 3 [HERE](#) and data description [HERE](#)