COSC 6323 – Statistical Methods in Research

Homework 1

January 26, 2018

Problems

1. The Excel file "Criminal Justice Expenditures" gives data on population (in thousands) and expenditures on criminal justice activities (in millions of dollars) for the 50 states and the District of Columbia as obtained from the 2005 Statistical Abstract of the United States.
   a. Describe the distribution of states' criminal justice expenditures with whatever measures appear appropriate. Comment on the features and implications of these data.
   b. Compute the per capita expenditures (EXPEND/POP) for these data. Repeat part (a). Discuss any difference in the nature of the distribution you may have stated in part (a).
   c. Make a scatterplot of total and per capita expenditures on the vertical axis against population on the horizontal axis. Which of these plots is more useful?

2. The data set in Excel file "Down Syndrome-Australia" lists all cases of Down syndrome in Victoria, Australia, from 1942 through 1957, as well as the number of births classified by the age of the mother (Andrews and Herzberg, 1985).
   a. Construct a relative frequency histogram for total number of births by age group.
   b. Construct a relative frequency histogram for number of mothers of Down syndrome patients by age group.
   c. Compare the shape of the two histograms. Does the shape of the histogram for Down syndrome suggest that age alone accounts for number of Down syndrome patients born?
   d. Construct a scatter diagram of total number of births versus number of mothers of Down syndrome. Does the scatter diagram support the conclusion in part (c)?

3. A lottery that sells 150,000 tickets has the following prize structure:
   a. First prize of $50,000
   b. 5 second prizes of $10,000
   c. 25 third prizes of $1000
   d. 1000 fourth prizes of $10
      i. Let Y be the winning amount of a randomly drawn lottery ticket. Describe the probability distribution of Y.
      ii. Compute the mean expected value of the ticket.

UNIVERSITY of
HOUSTON

iii. If the ticket costs $1.00, is the purchase of the ticket worthwhile? Explain your answer.

iv. Compute the standard deviation of this distribution. Comment on the usefulness of the standard deviation as a measure of dispersion for this distribution.

4. A teacher wants to curve her grades such that 10% are below 60 and 10% above 90. Assuming a normal distribution, what values of μ and $\sigma2$ will provide such a curve.

5. The following pose conceptual hypothesis situations. For each situation define H_0 and H_1 so as to provide control of the more serious error. Justify your choice and comment on logical values for α.

   a. You are deciding whether you should take an umbrella to work.
   b. You are planning a proficiency testing procedure to determine whether some employees should be fired.
   c. Same as part (b) except you want to determine whether some employees deserve a special merit raise.
   d. A cigarette manufacturer is conducting a test of nicotine content in order to justify a new advertising claim.
   e. You are considering the procedure to decide guilt or innocence in a court of law.
   f. You are considering whether you should buy a new battery for your calculator before the next statistics test.
   g. As a university administrator you are considering a policy to restrict student driving inn order to improve scholastic achievement.

6. An experiment is conducted to determine whether a new computer program will speed up the processing of credit card billing at a large bank. The mean time to process billing using the current program is 12.3 min with a standard deviation of 3.5 min. The new program is tested with 100 billings and yielded a sample mean of 10.9 min. Assuming the standard deviation of times in the new program is the same as the old, does the new program significantly reduce the time of processing? Use $\alpha=0.05$.

Submission

- Zip your code/documents and submit to blackboard. Name your zip file "**FirstName_LastName_HW1.zip**". One submission per individual.
- The homework should be submitted on BlackBoard by **11:59pm on February 09, 2018**.
- The homework report should be a pdf generated in Latex with figures from R only. All the figures and plots should have appropriate labels, titles and annotations.
- The code should also be provided.

Grading

Total points **100 pts** (1-2 take 10 pts each; 3-6 take 20 pts each)

UNIVERSITY of
HOUSTON