Problem Set 3

M. Leeney

March 10, 2016

1 Intervals and Hypotheses

Q1. In order to test the durability of as new paint, a highway department has test strips painted across heavily traveled roads in eight different locations. If, on the average, the test strips disappear after they had been crossed by 168,479 cars and the standard deviation is 12,851 cars construct a .99 confidence interval for the number of cars it will actually take on average to wear off the paint.Also determine a .95 confidence interval for the standard deviation of the population.

Q2. A sample survey of $N = 1000$ people with private health insurance (with a specific insurance provider) shows that $X = 42$ people have lied on their insurance application forms (about one or more issues). Construct a .95 confidence interval for the true proportion of those insured who have lied on their application forms. You may assume that for large samples (N large)that the quantity

$$Z = \frac{X - N\theta}{\sqrt{N\theta(1-\theta)}}$$

has a standard normal distribution, where θ is the population proportion. Comment. Note that $\theta(1-\theta) \leq 1/4$. Prove this fact!

Q3.Given the data in the table on sheet titled "TeamATeamB" in the public BigData folder test the hypothesis $H_0 =$ Population A has the same mean as population B against the alternative $H_1 =$ Mean of population A is greater than mean of population B

Q4. Given the paired data below determine the least squares line that fits the data. Compute the correlation coefficient and comment on the result. Can you find a curve that fits the data better? Sketch the data.

$$\begin{pmatrix} 1.2 & , & 1.5 \\ 1.7 & , & 1.9 \\ 2.3 & , & 2.5 \\ 2.5 & , & 2.8 \\ 3.1 & , & 3.8 \\ 3.4 & , & 4.4 \\ 4.2 & , & 6.5 \\ 4.3 & , & 6.9 \end{pmatrix}$$

Q5. Briefly discuss the statement "Complete disorder is an impossibility" - max 2-handwritten A4 pages.

Q5. Calculate the eigenvalues and eigenvectors of the following matrix

$$\begin{pmatrix} 3 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}$$

Q6. Given the data points $(4,3),(2,3),(1,1),(-3,-1),(-2,-4),(-2,-2)$, calculate the Covariance matrix - calculate its eigenvalues and the corresponding eigen directions.

Q7. Given the data points $(3,2),(2,2),(1,1),(0,0),(-2,2),(-1,-2),(-1,-1)$ trace the 2-means calculation for three iterations approximating the movement of the means and the corresponding clusters. Start at any two points you wish.