Dr. Özlem İLK DAĞ

Spring 2013-2014

**STAT 552**

**HOMEWORK 1**

**Due 25 March 2014, Tuesday, 14:40**

You should work on these questions on your own. Please feel free to get help from me, but not from anyone else.

1. Computer application: We will understand the central limit theorem better by two examples and via following these steps:

***Example 1:*** *Step I:* Assume that you only have random numbers from continuous Uniform distribution between 0 and 1. Provide the **theory** on how you can generate random numbers from the following probability density function:



**Hint:** Remember the use of probability integral transformation method from S551 for random number generation.

*Step II:* Use a software (e.g. R or MATLAB etc.) to generate 50 random numbers from this distribution with a=1 and b=5, and to calculate the sample mean.

*Step III:* Repeat step II for 1,000 times (i.e., generating random numbers and calculating means). Save these 1,000 means so that you can use them in the next step.

*Step IV:* Draw a histogram and a Q-Q plot of the 1,000 sample means. Apply a test for normality such as Shapiro-Wilk test on these 1,000 sample means.

Include the theory, your software code (if you used R, MATLAB etc.), graphs and test result (obtained in Step IV) in your answer.

What did you observe at the end of this experiment?

The distribution used in this example is a continuous one with a shape similar to normal anyway. Let’s see what happens when we have a discrete data:

***Example 2:*** *Step I:* Generate n=10 random numbers from a Poisson distribution with mean=1. You can use readily available functions to do this (That is, you don’t need to work through probability integral transformation method or theory to generate these numbers). Calculate the sample mean.

*Step II:* Repeat step I for 1,000 times (i.e., generating random numbers and calculating means). Save these 1,000 means so that you can use them in the next steps. Save (any) one of the data sets you generated (You don’t need to save all 1,000 samples, just one is enough here).

*Step III:* Draw a histogram and a Q-Q plot of the one sample. Apply a test for normality such as Shapiro-Wilk test on this one sample.

*Step IV:* Draw a histogram and a Q-Q plot of the 1,000 sample means. Apply a test for normality such as Shapiro-Wilk test on these 1,000 sample means.

*Step V:* Repeat the same experiment, now with n=100.

Include the theory, your software code (if you used R, MATLAB etc.), graphs and test results (obtained in Steps III, IV and V) in your answer.

What did you observe at the end of this experiment?

1. Let X denote the proportion of allocated time that a randomly selected student spends working on a certain aptitude test. Suppose that the pdf of X is

 where θ > -1. A random sample of ten students yields the following data: 0.92, 0.79, 0.9,0.65,0.86,0.47,0.73,0.97,0.94,0.77.

1. Show that this is a valid pdf.
2. Is this a member of regular exponential family? Why / why not?
3. Use the method of moments to obtain an estimator of θ, and then compute the estimate for this data.
4. Obtain the maximum likelihood estimator of θ, and then compute the estimate for the given data.
5. Consider θ in the range of [-1, 6] increasing by 1 unit (that is, consider θ={-1,0,1,2,…,6}). Plot θ versus likelihood values at each of the θ value. Also, plot θ versus *log-*likelihood values at each of the θ value. You can do this in any of the software you would like. Can you guess the MLE from these plots? Does it coincide with what you found in part d)?