PESIT Department of Computer Science and Engineering

Course: Data Analysis
Semester: 2014 Spring (January – May)
Instructor: BNR (Dr. B. Narsing Rao)

Assignment: 06
Topic: Confidence Intervals
Due by: See below
Method: Send zip archive (.zip, .rar, etc.) by email to bnrao@pes.edu
The name of the zip archive should be: DA-A06-your USN-your name
(USN must be upper case and your name should be in mixed case)
The zip archive should contain the following (see below for details):

1. PDF report including the output (see below) along with your observations
2. Source file containing R functions (named DA-A06-USN-Name.R) used

Part A

Due by: Midnight on Tuesday, March 4, 2014
Method: Send a pdf file by email to bnrao@pes.edu
The name of the file should be: DA-A06-USN-Name

Use the file ban-data.csv that was supplied with Assignment A-04. For this problem use the R functions qnorm etc. as necessary. For this assignment we will use only the data on income which will be considered to be the **population**.

- Generate, using R, random samples of size 30
- Compute the mean income in the sample
- Compute the 95% confidence interval using the **population** standard deviation (in this case, known)
- Determine whether the population mean lies within this confidence interval
- Repeat the experiment for a different number of samples
- Repeat the entire experiment using samples of size 60
- Produce reports showing the results (for each sample size) as in the following table:

Sample size:

Width of 95% confidence interval:

Number of samples	Proportion of samples where the population mean lay within the confidence interval around the sample mean
100	
1,000	
10,000	

If you feel more enthusiastic, you can show a plot such as the one in Figure 5.4 on page 329 of the textbook for the 100 sample case. Your pdf file should contain the tables, optional plot, and your observations on the results of the experiment.

Part B

a) Problem No. 14 on page 343 of the textbook
b) The following link explains the survey methodology used for the India survey:

http://www.pewglobal.org/2014/02/26/indians-want-political-change/2/#india-survey-methods

The following statement appears on this page: “The margin of sampling error is ±3.8 percentage points.”

Explain what this applies to (i.e. how it should be used to interpret the results of their poll), and how it was calculated.

Submit handwritten solution in class at 1:25 P.M. on Wednesday, March 5, 2014