**Laboratory Exercise Week 7**

Name- Virender Singh and Section- 003, 10 pts

9th October, 2020

*Directions*:

* Write your R code inside the code chunks after each question.
* Write your answer comments after the # sign.
* To generate the word document output, click the button Knit and wait for the word document to appear.
* RStudio will prompt you (only once) to install the knitr package.
* Submit your completed laboratory exercise using Blackboard’s Turnitin feature. Your Turnitin upload link is found on your Blackboard Course shell under the Laboratory folder.

For this exercise, you will need to use the package mosaic to find numerical and graphical summaries.

*# install package if necessary*  
**if** (**!require**(mosaic)) **install.packages**("mosaic")  
*# load the package in R*  
**library**(mosaic) *# load the package mosaic to use its functions*

* Consider the population of CEO salaries in the lesson this week.
* Select a random sample of 40 CEO’s and compute the sample mean salary and sample standard deviation salary.
* Use the do() function to compute the sample mean of 100 randomly selected samples. Plot the histogram of these 100 sample means.
* Compute the average and standard deviation of these 100 sample means. Are the values what you expect from the properties of the sampling distribution.

**Code chunk**

*# start your code*  
ceo\_salary <- **read.csv**("https://www.siue.edu/~jpailde/CEO\_Salary\_2012.csv")  
(samp1e <- **sample\_n**(ceo\_salary, size = 40))

## ï..Rank Name Company  
## 1 215 Michael T Strianese L-3 Communications Holdings  
## 2 412 R Adam Norwitt Amphenol  
## 3 57 Scott D Sheffield Pioneer Natural Resources  
## 4 264 Enrique T Salem Symantec  
## 5 239 James J Volker Whiting Petroleum  
## 6 300 Francis S Blake Home Depot  
## 7 195 James W Griffith Timken  
## 8 477 Martin Mucci Paychex  
## 9 67 Ian M Cook Colgate-Palmolive  
## 10 399 Paul C Reilly Raymond James Financial  
## 11 392 Marc N Casper Thermo Fisher Scientific  
## 12 47 David B Snow Jr Medco Health Solutions  
## 13 479 Steven A Ballmer Microsoft  
## 14 393 Christopher M Crane Exelon  
## 15 284 Thomas C Gallagher Genuine Parts  
## 16 379 Inge G Thulin 3M  
## 17 172 David W Nelms Discover Financial Services  
## 18 385 Hikmet Ersek Western Union  
## 19 469 Beth E Mooney KeyCorp  
## 20 42 K Rupert Murdoch News Corp  
## 21 37 Mark Donegan Precision Castparts  
## 22 268 Jack B Moore Cameron International  
## 23 152 James E Rohr PNC Financial Services Group  
## 24 278 James P Rogers Eastman Chemical  
## 25 247 Brian Duperreault Marsh & McLennan Cos  
## 26 316 Gregory L Henslee O'Reilly Automotive  
## 27 139 Surya N Mohapatra Quest Diagnostics  
## 28 282 Benjamin G S Fowke III Xcel Energy  
## 29 261 Michael J Long Arrow Electronics  
## 30 454 Edward J Bonach CNO Financial  
## 31 446 Gary C Kelly Southwest Airlines  
## 32 475 Alan S Armstrong Williams Cos  
## 33 415 D Mark Durcan Micron Technology  
## 34 56 Frank R Martire Fidelity National Information Services  
## 35 370 Ursula M Burns Xerox  
## 36 102 J Wayne Leonard Entergy  
## 37 101 Louis R Chenevert United Technologies  
## 38 183 Debra A Cafaro Ventas  
## 39 142 Sandeep Mathrani General Growth Properties  
## 40 289 Kenneth C Frazier Merck & Co  
## Annual.pay Age  
## 1 8.160 56  
## 2 3.075 43  
## 3 20.050 59  
## 4 6.560 46  
## 5 7.290 65  
## 6 5.675 62  
## 7 8.890 58  
## 8 1.450 52  
## 9 19.235 60  
## 10 3.310 57  
## 11 3.445 44  
## 12 22.195 57  
## 13 1.380 56  
## 14 3.426 53  
## 15 6.050 64  
## 16 3.746 58  
## 17 9.880 51  
## 18 3.540 52  
## 19 1.695 56  
## 20 24.790 81  
## 21 25.620 54  
## 22 6.460 58  
## 23 11.140 63  
## 24 6.120 61  
## 25 7.045 64  
## 26 5.080 51  
## 27 12.575 62  
## 28 6.066 54  
## 29 6.590 52  
## 30 2.076 58  
## 31 2.290 57  
## 32 1.485 50  
## 33 2.986 51  
## 34 20.785 63  
## 35 4.025 53  
## 36 15.110 61  
## 37 15.330 54  
## 38 9.495 54  
## 39 11.980 50  
## 40 5.935 57

(mean.sal <- **mean**(**~** Annual.pay, data = samp1e))

## [1] 8.550875

(sd.sal <- **sd**(**~** Annual.pay, data = samp1e))

## [1] 6.821359

means <- **do**(100)**\*mean**(**~** Annual.pay,  
 data = **sample\_n**(ceo\_salary, size = 40))  
**histogram**(**~** mean, data = means,   
 main = "Sample Mean Salary Distribution",   
 xlab = "Sample Mean Salary ($Mil) of 40 CEO's")



**sapply**(means, mean)

## mean   
## 10.64072

**sapply**(means, sd)

## mean   
## 1.691713

*# These values were expected as the mean must be center of the distribution which is approximately 10 and we also have a low standard deviation indicating only a few outliers.*  
*# last R code line*

* Rockwell hardness of pins of a certain type is known to have a mean value of 50 and standard deviation of 1.2.
* If the distribution is normal, what is the probability that the sample mean hardness for a random sample of 9 pins is at least 51?
* Without assuming population normality, what is the (approximate) probability that the sample mean hardness for a random sample of 40 pins is at least 51?

**Code chunk**

*# start your code*  
1**-pnorm**(q = 51, mean = 50, sd = 1.2**/sqrt**(9))

## [1] 0.006209665

1**-pnorm**(q = 51, mean = 50, sd = 1.2**/sqrt**(40))

## [1] 6.804011e-08

*# last R code line*

* Suppose that a random sample of size 64 is to be selected from a population with mean 40 and standard deviation 5.
* What are the mean and standard deviation of the sampling distribution of the sample mean?
* What is the approximate probability that the sample mean will be within 0.5 of the population mean?

**Code chunk**

*# start your code*  
randsamp <- **rnorm**(n = 64, mean = 40, sd = 5)  
**mean**(randsamp)

## [1] 40.52327

**sd**(randsamp)

## [1] 4.49331

**pnorm**(q = **mean**(randsamp)**+**0.5, mean = 40, sd = 5**/sqrt**(64)) **-** **pnorm**(q = **mean**(randsamp)**-**0.5, mean = 40, sd = 5**/sqrt**(64))

## [1] 0.434359

*# last R code line*