**Module Assignment**

**Module 2**

**QMB-6304 Foundations of Business Statistics**

**#Debjani Sarma**

**rm(list=ls())**

**library(rio)**

**library(moments)**

**getwd()**

Write a simple R script to execute the following:

**Preprocessing:**

Load the data in “6304 Assignment 2 Data.xlsx” into an object. The file includes 10000 observations for each of four variables, which are creatively named for the four Marx Brothers.

**marxbrothers=import("6304 Module 2 Assignment Data.xlsx")**

**colnames(marxbrothers)=tolower(make.names(colnames(marxbrothers)))**

**attach(marxbrothers)**

**Analysis:**

1. Use common tools to determine whether any of the four variables are normally distributed. Explain how you arrived at your conclusions.

**> skewness(harpo)**

**[1] 0.5786532**

**> skewness(zeppo)**

**[1] -0.0245893**

**> skewness(groucho)**

**[1] -0.01466599**

**> skewness(chico)**

**[1] -2.031371**

**> kurtosis(harpo)**

**[1] 3.364361**

**> kurtosis(zeppo)**

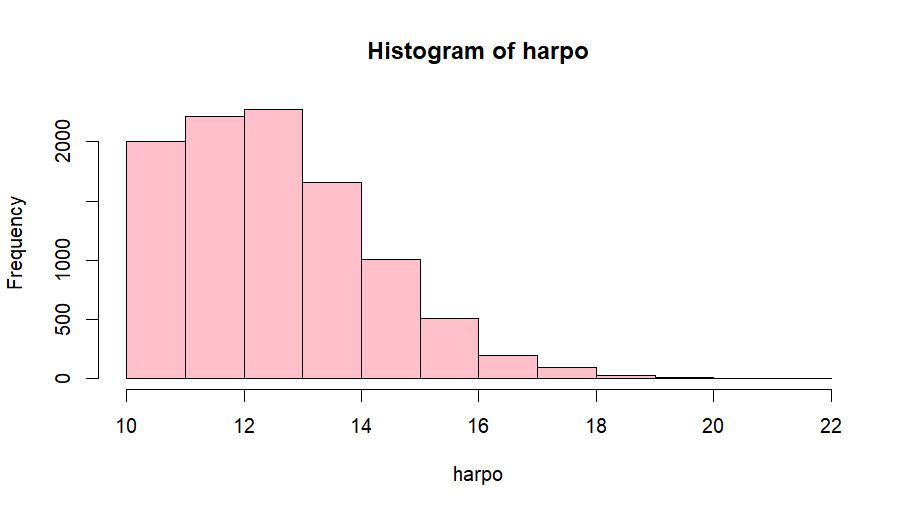
**[1] 3.019482**

**> kurtosis(groucho)**

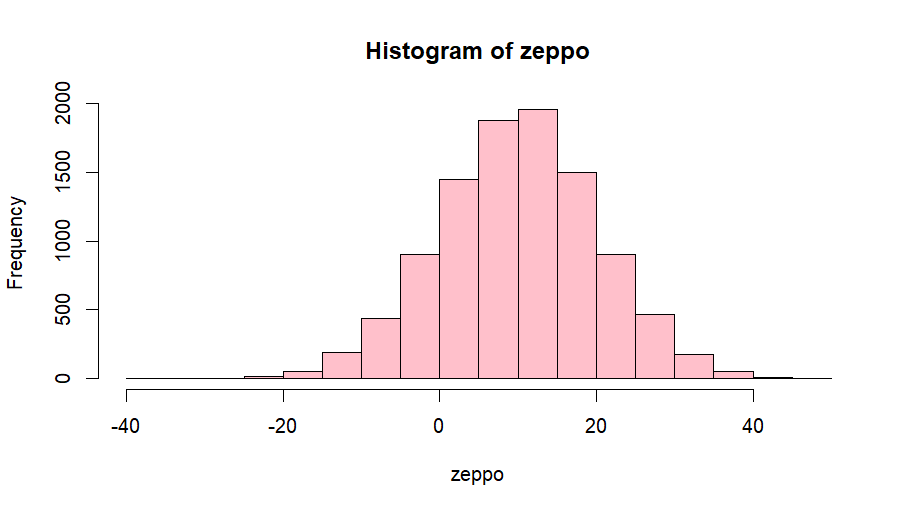
**[1] 1.795703**

**> kurtosis(chico)**

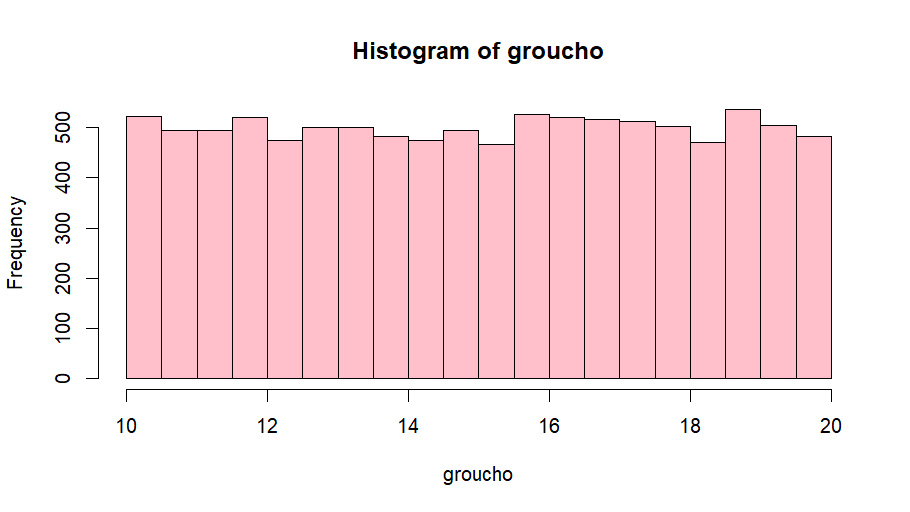
**[1] 9.268749**

**> hist(harpo,col="pink",main="Histogram of harpo")**

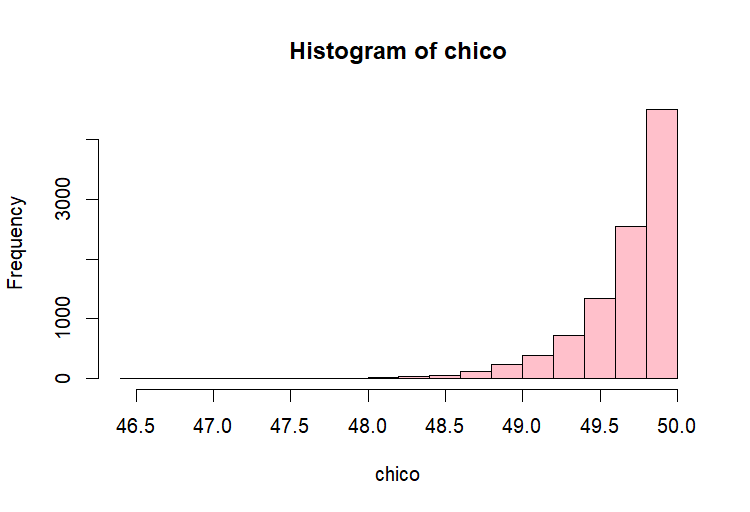
**> hist(zeppo,col="pink",main="Histogram of zeppo")**



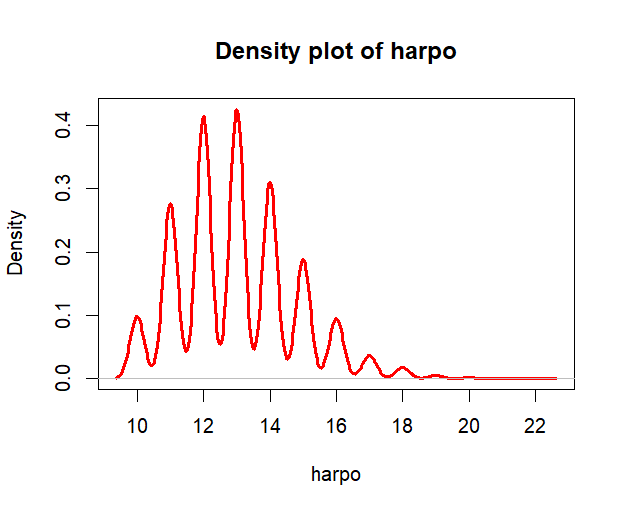
**> hist(groucho,col="pink",main="Histogram of groucho")**

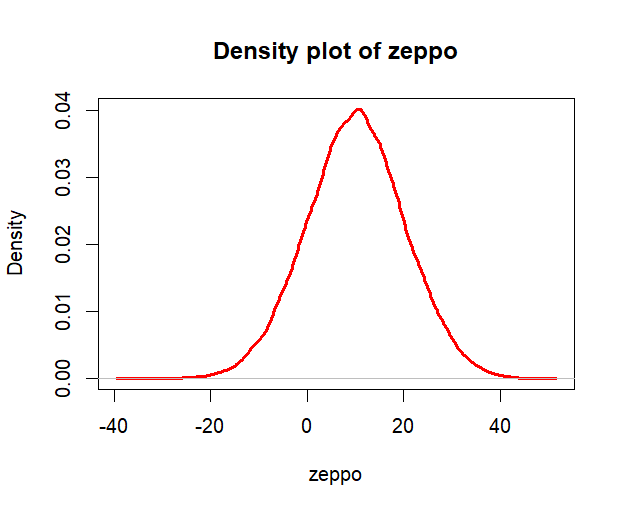


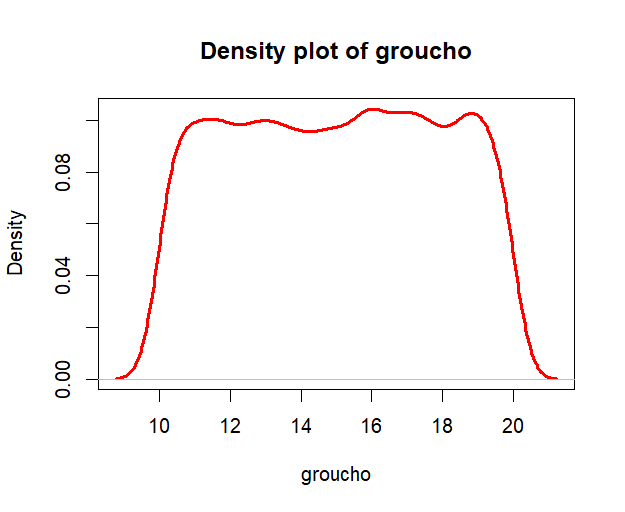
**> hist(chico,col="pink",main="Histogram of chico")**



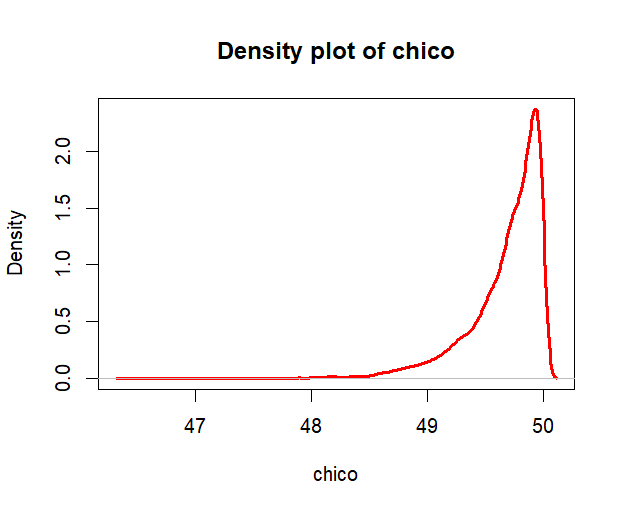
**> plot(density(harpo),lwd=3,col="red",xlab="harpo",main="Density plot of harpo")**



**> plot(density(zeppo),lwd=3,col="red",xlab="zeppo",main="Density plot of zeppo")**

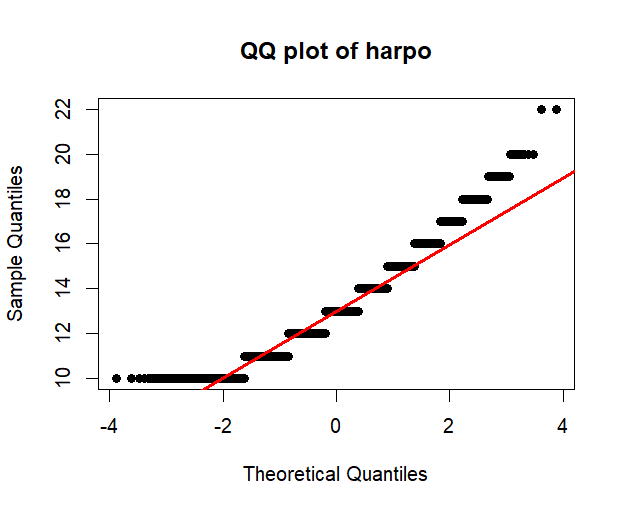
**>plot(density(groucho),lwd=3,col="red",xlab="groucho",main="Density plot of groucho")**

**> plot(density(chico),lwd=3,col="red",xlab="chico",main="Density plot of chico")**



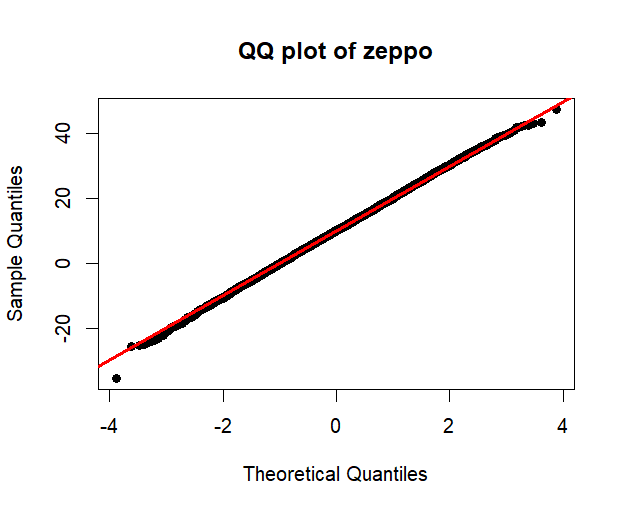
**> qqnorm(harpo,pch=19,main="QQ plot of harpo")**

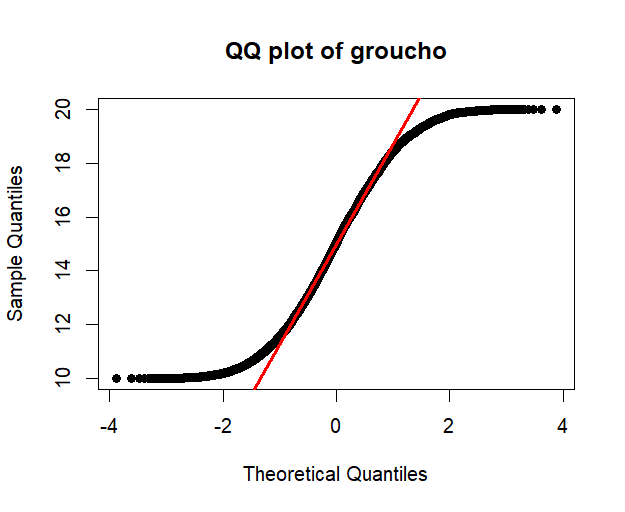
**> qqline(harpo,col="red",lwd=3)**



**> qqnorm(zeppo,pch=19,main="QQ plot of zeppo")**

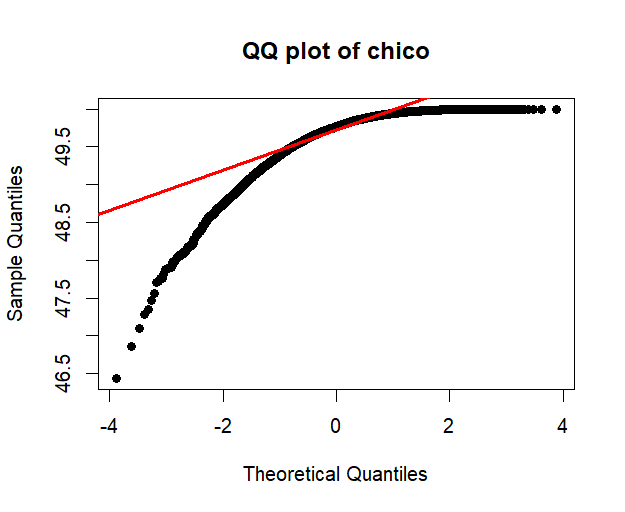
**> qqline(zeppo,col="red",lwd=3)**



**> qqnorm(groucho,pch=19,main="QQ plot of groucho")**

**> qqline(groucho,col="red",lwd=3)**

**> qqnorm(chico,pch=19,main="QQ plot of chico")**

**> qqline(chico,col="red",lwd=3)**

**> mean.zeppo=mean(zeppo)**

**> mean.zeppo**

**[1] 10.06579**

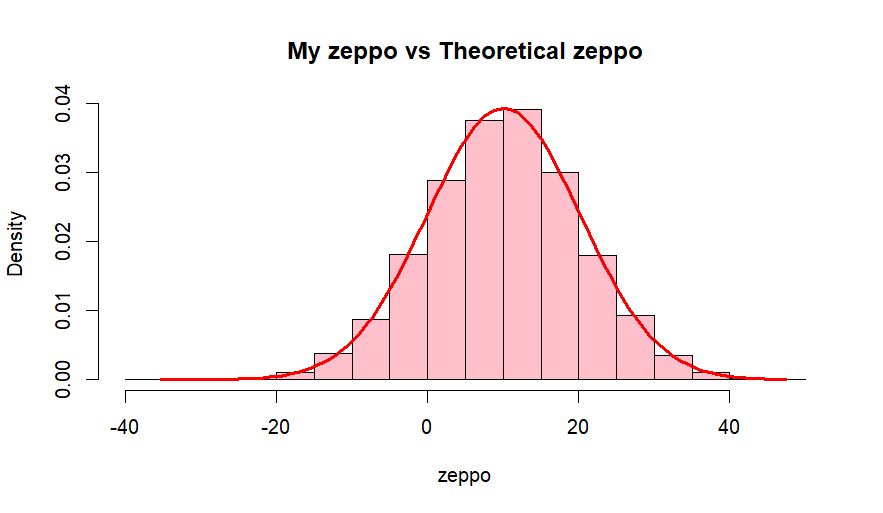
**> sd.zeppo=sd(zeppo)**

**> sd.zeppo**

**[1] 10.14926**

**> hist(zeppo,col="pink",main="My zeppo vs Theoretical zeppo",probability=TRUE)**

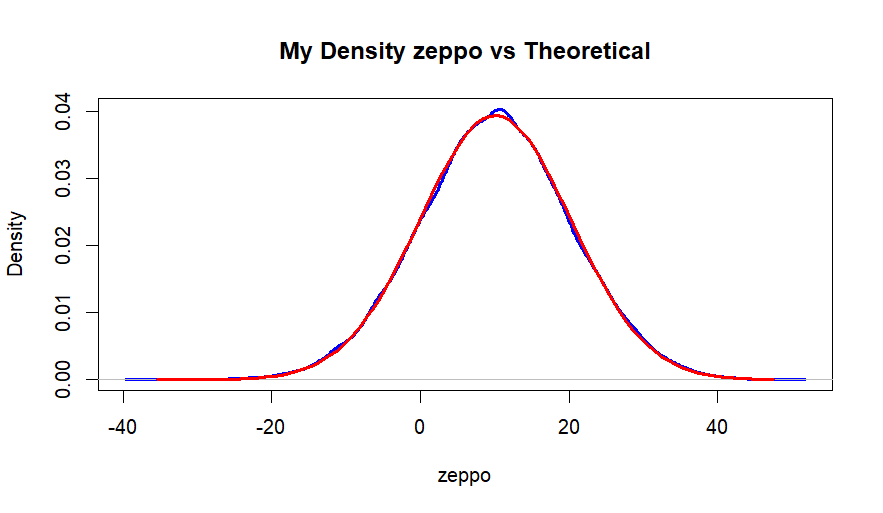
**> curve(dnorm(x,mean.zeppo, sd.zeppo), from=min(zeppo), to = max(zeppo), col="red",lwd=3,add=TRUE)**



**> plot(density(zeppo),lwd=3,col="blue",xlab="zeppo",main="My Density zeppo vs Theoretical")**

**> curve(dnorm(x,mean.zeppo, sd.zeppo), from=min(zeppo), to = max(zeppo), col="red",lwd=3,add=TRUE)**

**>**



By using common tools of normal distribution such as skewness, kurtosis, histogram, density plot, QQplot, we can conclude that zeppo is normally distributed. Zeppo has a skewness =

-0.0245893 and kurtosis = 3.019482.We can also see the difference between histogram of sample zeppo vs theoretical zeppo, as well as density plot of sample zeppo vs theoretical zeppo,both of which tends to follow a normal distribution.

1. Focus on the Chico variable. Build a sampling distribution of the population mean by calculating the mean for each of 1000 samples of n=40. Plot the sampling distribution with an appropriate graph title. Verify whether these 1000 means are in fact normally distributed and justify your conclusion with appropriate analytics and/or graphical tools.

**sample\_mean=c()**

**for (i in 1:1000) {**

**chico.sample=marxbrothers[sample(1:nrow(marxbrothers),40),]**

**mean.chico=mean(chico.sample$chico)**

**sample\_mean=c(sample\_mean,mean.chico)**

**}**

**sample\_mean**

**> sample\_mean**

**[1] 49.76433 49.78129 49.66358 49.66766 49.65651 49.65053 49.73657 49.66465 49.63075 49.70403**

**[11] 49.72424 49.70720 49.69205 49.64220 49.67486 49.81837 49.70835 49.56490 49.73638 49.71341**

**[21] 49.76720 49.66704 49.57992 49.67786 49.65831 49.65371 49.67789 49.58598 49.64506 49.71128**

**[31] 49.53677 49.72844 49.69730 49.64624 49.66912 49.70653 49.69557 49.70772 49.64479 49.73742**

**[41] 49.74959 49.63843 49.74407 49.60182 49.66309 49.65217 49.69559 49.70839 49.67359 49.75151**

**[51] 49.69119 49.74621 49.67901 49.68382 49.65778 49.65192 49.72474 49.67940 49.74259 49.66092**

**[61] 49.63842 49.58250 49.73897 49.58584 49.69278 49.67443 49.69845 49.73816 49.67205 49.66674**

**[71] 49.67432 49.69883 49.56200 49.69960 49.63409 49.64687 49.71115 49.70049 49.57132 49.63116**

**[81] 49.76740 49.72464 49.67387 49.61742 49.64766 49.59314 49.67765 49.68009 49.69934 49.66487**

**[91] 49.70757 49.65969 49.72051 49.67472 49.70310 49.57855 49.47923 49.65044 49.69125 49.65693**

**[101] 49.71428 49.63654 49.74086 49.70213 49.76260 49.66563 49.60195 49.61188 49.70263 49.73075**

**[111] 49.76906 49.71673 49.75681 49.64175 49.68577 49.61357 49.68890 49.67626 49.71309 49.72499**

**[121] 49.65043 49.65123 49.72387 49.66668 49.64468 49.64233 49.61697 49.62933 49.61061 49.71480**

**[131] 49.67470 49.58589 49.62053 49.73838 49.60335 49.72162 49.71118 49.63002 49.73111 49.72142**

**[141] 49.67469 49.64758 49.61998 49.68069 49.64835 49.66871 49.52000 49.61301 49.59302 49.72679**

**[151] 49.68559 49.65461 49.66440 49.65507 49.65753 49.60606 49.72258 49.72596 49.70674 49.69688**

**[161] 49.70297 49.65162 49.58428 49.60329 49.65274 49.60619 49.66888 49.62744 49.67754 49.68891**

**[171] 49.63941 49.63374 49.69172 49.70613 49.70001 49.72594 49.67291 49.63066 49.70151 49.64821**

**[181] 49.60104 49.57839 49.60484 49.68732 49.73435 49.60456 49.68553 49.65781 49.72134 49.57310**

**[191] 49.75653 49.62540 49.66503 49.71892 49.67939 49.63850 49.63274 49.70797 49.63006 49.75020**

**[201] 49.66405 49.63073 49.69974 49.59596 49.68912 49.60787 49.75289 49.68034 49.55313 49.62949**

**[211] 49.66726 49.65290 49.73264 49.78287 49.68934 49.69057 49.69713 49.74931 49.66892 49.64196**

**[221] 49.70026 49.62829 49.71518 49.66213 49.69697 49.74466 49.60658 49.66515 49.68822 49.66527**

**[231] 49.64630 49.77191 49.69836 49.67090 49.65348 49.61834 49.72880 49.70528 49.70233 49.73178**

**[241] 49.65011 49.66748 49.68099 49.59746 49.66835 49.63726 49.70044 49.65436 49.66359 49.56613**

**[251] 49.64437 49.73364 49.73203 49.66754 49.63423 49.69745 49.68695 49.63979 49.69697 49.68323**

**[261] 49.63205 49.60948 49.65443 49.67413 49.71656 49.59637 49.62473 49.69878 49.63300 49.72938**

**[271] 49.59816 49.78442 49.67762 49.65133 49.64352 49.77662 49.66299 49.70404 49.70649 49.72929**

**[281] 49.69429 49.59360 49.69377 49.72998 49.74748 49.69116 49.72088 49.67147 49.67573 49.68734**

**[291] 49.73992 49.68726 49.68994 49.72349 49.76094 49.63939 49.62910 49.68499 49.68638 49.67861**

**[301] 49.68909 49.70982 49.58412 49.73559 49.61726 49.65253 49.76199 49.64970 49.75897 49.74723**

**[311] 49.69962 49.65536 49.71598 49.71384 49.67259 49.65035 49.65349 49.58204 49.64111 49.71609**

**[321] 49.71591 49.63442 49.58202 49.54978 49.67112 49.66749 49.69162 49.67396 49.68494 49.75505**

**[331] 49.70031 49.68291 49.61952 49.70167 49.65123 49.63186 49.63465 49.66325 49.58594 49.70727**

**[341] 49.73761 49.74694 49.60474 49.66763 49.64068 49.67924 49.67980 49.70218 49.64098 49.70913**

**[351] 49.70964 49.68478 49.62398 49.68593 49.72799 49.66319 49.72345 49.64359 49.68103 49.67560**

**[361] 49.68357 49.71481 49.68803 49.73117 49.58127 49.67201 49.69924 49.59735 49.61240 49.65432**

**[371] 49.63385 49.72073 49.68338 49.74122 49.71794 49.72088 49.58947 49.71394 49.69787 49.57919**

**[381] 49.68952 49.66353 49.68847 49.62999 49.67872 49.71396 49.71800 49.66332 49.64665 49.71681**

**[391] 49.65957 49.66873 49.68723 49.68678 49.68652 49.64445 49.56754 49.74227 49.70523 49.64788**

**[401] 49.63144 49.67459 49.71186 49.63965 49.76252 49.70242 49.71351 49.71322 49.60668 49.59592**

**[411] 49.72062 49.62746 49.67991 49.67258 49.64099 49.60788 49.60478 49.73176 49.70725 49.65562**

**[421] 49.66834 49.72616 49.66280 49.69658 49.63473 49.71591 49.58388 49.67216 49.61105 49.76373**

**[431] 49.66669 49.72555 49.58783 49.57708 49.73364 49.69519 49.61252 49.65426 49.71249 49.73532**

**[441] 49.76573 49.78267 49.63903 49.53663 49.61475 49.64823 49.67011 49.75081 49.64356 49.59860**

**[451] 49.63114 49.59540 49.63689 49.65004 49.66123 49.67203 49.59221 49.61345 49.65927 49.61736**

**[461] 49.70529 49.58708 49.63933 49.76524 49.73593 49.68952 49.61425 49.60111 49.72104 49.73137**

**[471] 49.75034 49.67782 49.66283 49.71813 49.69362 49.66191 49.68796 49.72738 49.65611 49.69678**

**[481] 49.73118 49.63547 49.74587 49.67212 49.66984 49.65365 49.68868 49.74836 49.63950 49.71921**

**[491] 49.66307 49.70852 49.71046 49.62649 49.67751 49.66662 49.71014 49.65902 49.70896 49.60945**

**[501] 49.61612 49.49634 49.66097 49.69030 49.70813 49.74327 49.65394 49.60531 49.66892 49.66176**

**[511] 49.70856 49.64611 49.66464 49.63847 49.71683 49.62952 49.63599 49.66645 49.71828 49.67016**

**[521] 49.71226 49.74314 49.70154 49.73042 49.65738 49.65811 49.59314 49.64441 49.70756 49.69523**

**[531] 49.64980 49.57097 49.65968 49.76257 49.67527 49.70399 49.67548 49.70802 49.68231 49.69896**

**[541] 49.72870 49.73862 49.71838 49.72024 49.66328 49.69056 49.61497 49.64282 49.72792 49.62570**

**[551] 49.74674 49.69066 49.70003 49.64411 49.71075 49.68114 49.61826 49.67617 49.71285 49.61201**

**[561] 49.70066 49.63239 49.57775 49.64295 49.75286 49.67495 49.70806 49.68143 49.72251 49.60617**

**[571] 49.70539 49.76651 49.66887 49.71387 49.65254 49.68514 49.72486 49.65753 49.63772 49.65075**

**[581] 49.74628 49.68943 49.61803 49.72910 49.60270 49.70781 49.71894 49.70498 49.63364 49.67151**

**[591] 49.68076 49.75969 49.63737 49.67202 49.68543 49.54919 49.63121 49.68367 49.69801 49.58357**

**[601] 49.62488 49.64392 49.72103 49.71602 49.65962 49.64299 49.78748 49.72763 49.56488 49.62780**

**[611] 49.67639 49.69974 49.62287 49.69431 49.64283 49.66997 49.64102 49.64344 49.55229 49.62554**

**[621] 49.69907 49.67763 49.63662 49.69462 49.67549 49.56746 49.70850 49.68301 49.64673 49.74383**

**[631] 49.67603 49.65748 49.70226 49.70594 49.76941 49.68591 49.66386 49.72753 49.75284 49.61538**

**[641] 49.59663 49.63513 49.70256 49.65654 49.67340 49.65479 49.72189 49.67181 49.62322 49.67934**

**[651] 49.61628 49.74902 49.65584 49.56348 49.63972 49.70934 49.76975 49.61175 49.69216 49.66875**

**[661] 49.71253 49.68516 49.68691 49.62464 49.69207 49.62628 49.61520 49.66831 49.69661 49.61061**

**[671] 49.69500 49.67809 49.70026 49.63854 49.60972 49.71531 49.68275 49.72027 49.70603 49.72285**

**[681] 49.64631 49.67751 49.67948 49.68020 49.70582 49.69700 49.70539 49.72367 49.68883 49.74371**

**[691] 49.64110 49.65609 49.64830 49.59920 49.67481 49.73313 49.65434 49.65908 49.72990 49.55884**

**[701] 49.61926 49.60645 49.59630 49.71170 49.70506 49.75119 49.69388 49.70149 49.64132 49.71322**

**[711] 49.59712 49.59391 49.63820 49.72194 49.60594 49.79634 49.72964 49.71382 49.64808 49.71322**

**[721] 49.79337 49.71123 49.67832 49.70732 49.61846 49.66134 49.65908 49.62921 49.64617 49.70610**

**[731] 49.71477 49.72307 49.74710 49.64677 49.64452 49.62263 49.77272 49.71735 49.67513 49.63301**

**[741] 49.69909 49.70748 49.70249 49.70614 49.57520 49.68270 49.65333 49.61097 49.67015 49.59346**

**[751] 49.65703 49.81204 49.71208 49.78706 49.69059 49.67679 49.63186 49.69475 49.70909 49.73089**

**[761] 49.64551 49.70732 49.72534 49.73031 49.55238 49.67229 49.69807 49.69707 49.63432 49.69643**

**[771] 49.70612 49.63830 49.62758 49.68930 49.67721 49.71457 49.70476 49.68038 49.67146 49.70281**

**[781] 49.68075 49.59967 49.63965 49.68323 49.70900 49.66859 49.66806 49.75962 49.74816 49.69737**

**[791] 49.69385 49.74665 49.65763 49.62626 49.66816 49.75658 49.66516 49.61166 49.73732 49.65339**

**[801] 49.66042 49.72395 49.76530 49.68183 49.63191 49.63377 49.67736 49.72080 49.73200 49.71706**

**[811] 49.61952 49.67081 49.62610 49.71941 49.61809 49.60203 49.58720 49.69760 49.74874 49.64297**

**[821] 49.68475 49.63552 49.71283 49.66071 49.59062 49.73003 49.68283 49.77899 49.59839 49.75949**

**[831] 49.53869 49.68299 49.68091 49.56981 49.75778 49.62179 49.69672 49.72369 49.69440 49.68317**

**[841] 49.61693 49.65572 49.59329 49.65485 49.75034 49.73668 49.72550 49.72083 49.65098 49.68467**

**[851] 49.72417 49.65159 49.66085 49.58574 49.66765 49.59833 49.61317 49.66871 49.58998 49.70655**

**[861] 49.57186 49.71035 49.68014 49.61565 49.59307 49.65939 49.69525 49.71994 49.69728 49.65643**

**[871] 49.72302 49.70710 49.65913 49.71746 49.69747 49.74087 49.66418 49.56411 49.67359 49.72112**

**[881] 49.65120 49.67999 49.64191 49.73289 49.65807 49.66045 49.68639 49.60923 49.68524 49.56390**

**[891] 49.75560 49.65522 49.56928 49.72066 49.65345 49.54735 49.64346 49.75145 49.66754 49.60625**

**[901] 49.68050 49.59429 49.66499 49.66581 49.65985 49.74881 49.73546 49.61177 49.71347 49.65842**

**[911] 49.59433 49.59926 49.68637 49.57662 49.66749 49.62796 49.67907 49.65385 49.66518 49.73564**

**[921] 49.54810 49.71760 49.71205 49.62797 49.70422 49.73781 49.70576 49.52791 49.62351 49.69925**

**[931] 49.69725 49.67907 49.75672 49.76550 49.72387 49.72431 49.65899 49.45643 49.67877 49.69534**

**[941] 49.63014 49.66661 49.57945 49.63026 49.68193 49.70585 49.70761 49.63217 49.70502 49.71785**

**[951] 49.68782 49.63878 49.66121 49.73622 49.55440 49.72696 49.58812 49.79275 49.73168 49.61845**

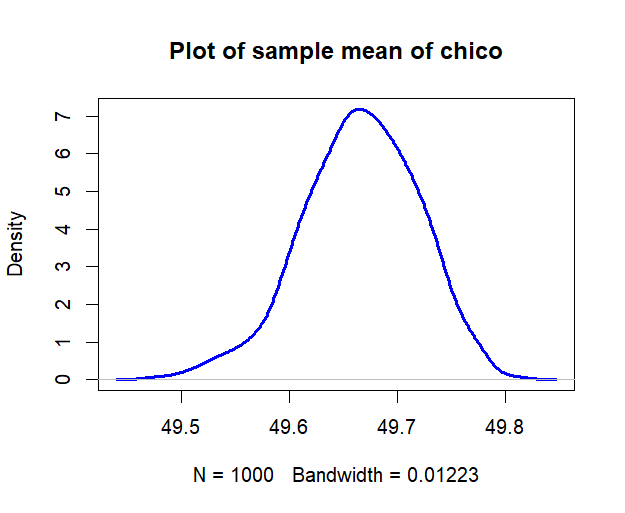
**[961] 49.68572 49.58742 49.63143 49.71913 49.77294 49.71616 49.69075 49.67826 49.62491 49.67831**

**[971] 49.59237 49.67045 49.60153 49.52459 49.66006 49.60812 49.58666 49.66699 49.73089 49.63120**

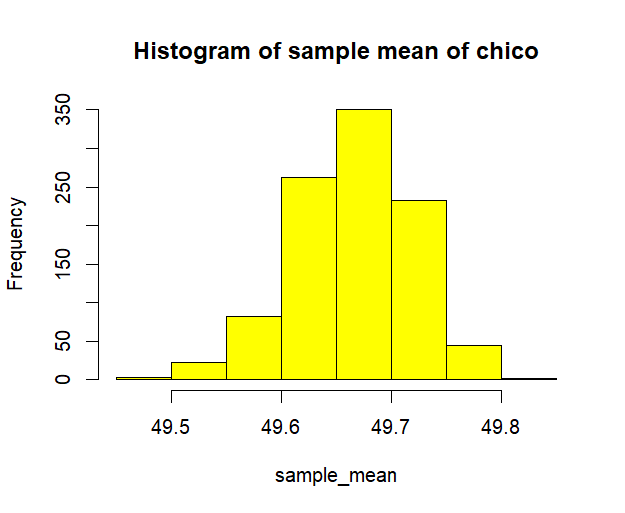
**[981] 49.66973 49.58786 49.74228 49.70079 49.67669 49.61151 49.65707 49.69979 49.70569 49.68353**

**[991] 49.71052 49.68740 49.71541 49.73542 49.71765 49.71802 49.69627 49.71882 49.68065 49.69076**

**> plot(density(sample\_mean),main="Plot of sample mean of chico",lwd=3,col="blue")**



**> hist(sample\_mean,col="yellow",main="Histogram of sample mean of chico")**



**> skewness(sample\_mean)**

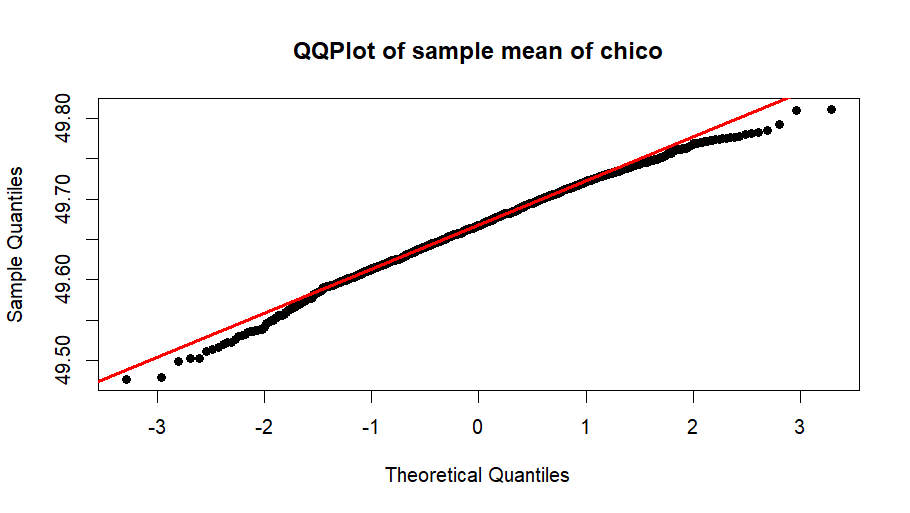
**[1] -0.3190146**

**> kurtosis(sample\_mean)**

**[1] 3.111668**

**> qqnorm(sample\_mean,main="QQPlot of sample mean of chico",pch=19)**

**> qqline(sample\_mean,lwd=3,col="red")**



**> sample.mean= mean(sample\_mean)**

**> sample.mean**

**[1] 49.66587**

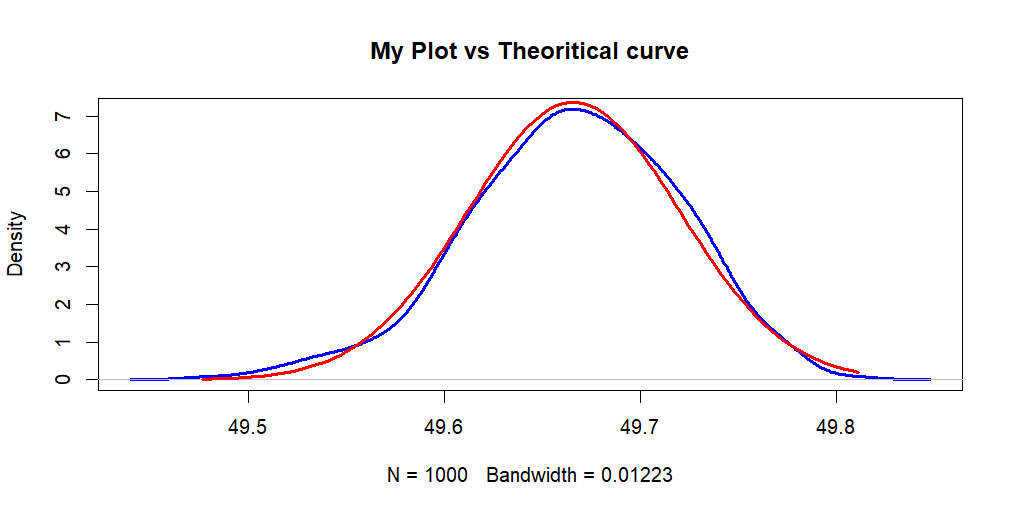
**> sd.mean = sd(sample\_mean)**

**> sd.mean**

**[1] 0.05409818**

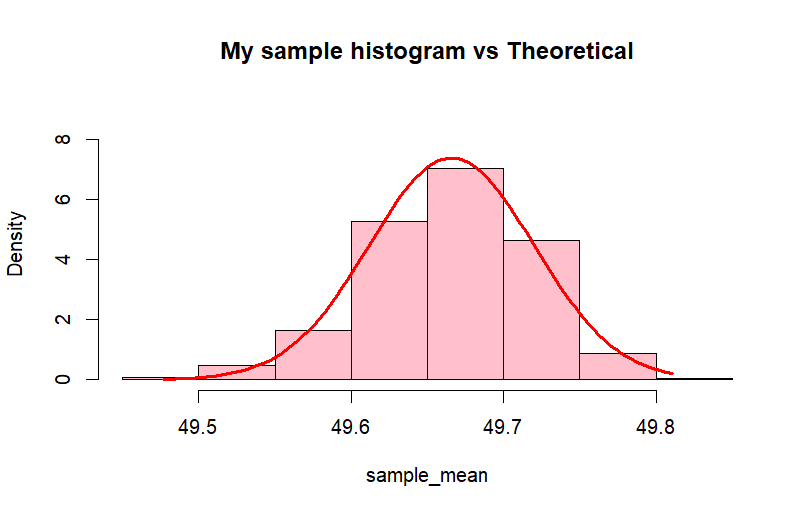
**> plot(density(sample\_mean),main="My Plot vs Theoritical curve",lwd=3,col="blue")**

**> curve(dnorm(x,sample.mean, sd.mean), from=min(sample\_mean), to = max(sample\_mean), col="red",lwd=3,add=TRUE)**



**> hist(sample\_mean,col="pink",main="My sample histogram vs Theoretical",probability=TRUE, ylim=c(0,9))**

**> curve(dnorm(x,sample.mean, sd.mean), from=min(sample\_mean), to = max(sample\_mean), col="red",lwd=3,add=TRUE)**



According to Central Limit Theorem, if we take large samples of a population, the mean of the sample is normally distributed even if the population is not normally distributed.

In this example, the population of chico is not normally distributed. Based on the QQplot, Density plot, Histogram, skewness and kurtosis of the sample mean of chico, we can conclude that sample mean of chico is normally distributed.

We can also see the difference between histogram of sample mean vs theoretical mean, as well as density plot of sample mean vs theoretical mean,both of which tends to follow a normal distribution.

Your deliverable will be a single MS-Word file showing 1) the R script which executes the above instructions and 2) the results of those instructions. The first line of your script file should be a “#” comment line showing your name as it appears in Canvas. Results should be presented in the order in which they are listed here. Deliverable due time will be announced in class and on Canvas. **This is an individual assignment to be completed before you leave the classroom. No collaboration of any sort is allowed on this assignment.**