R.M.K. COLLEGE OF ENGINEERING AND TECHNOLOGY

RSM Nagar, Puduvoyal– 601206, Gummidipoondi (T.K), Thiruvallur (D.T), Tamil Nadu

(An Autonomous Institution)

Approved by AICTE, New Delhi/ Affiliated to Anna University, Chennai

Accredited by NBA (All Eligible Courses) / NAAC with "A" GRADE

An ISO 21001:2018 Certified Institution

DEPARTMENT OF SCIENCE AND HUMANITIES

(MATHEMATICS)

Regulation2022

22MA401 – PROBABILITY AND STATISTICS

LABORATORY MANUAL

(Common to CSE & ADS)

Name :--

Register Number :---

Branch/Section :---

Academic Year/Semester :---

Course Code	**PROBABILITY AND STATISTICS** (Common to CSE, IT, CSD, AD)	L	T	P	C
22MA401	**SEMESTER IV** **(Theory Course with Laboratory Component)**	3	0	2	4

OBJECTIVES:

The course is designed to:

- Provide the necessary basic concepts of random variables and to introduce some standard distributions.
- Test the hypothesis for small and large samples.
- Introduce the concepts of analysis of variances.
- Understand the concept of statistical quality control**.**

UNIT I ONE-DIMENSIONAL RANDOM VARIABLES 9+6

Basic probability definitions- Independent events- Conditional probability (revisit) - Random variable - Discrete and continuous random variables - Moments - Moment generating functions - Binomial, Poisson, Geometric, Uniform, Exponential and Normal distributions.

List of Exercise/Experiments using R Programming:

1. Finding conditional probability.
2. Finding mean, variance and standard deviation.

UNIT II TWO-DIMENSIONAL RANDOM VARIABLES 9+6

Joint distributions - Marginal and conditional distributions - Covariance - Correlation and linear regression - Transformation of random variables.

List of Exercise/Experiments using R Programming:

1. Finding marginal density functions for discrete random variables.
2. Calculating correlation and regression.

UNIT III TESTING OF HYPOTHESIS 9+6

Sampling distributions - Estimation of parameters - Statistical hypothesis - Large sample tests based on Normal distribution for single mean and difference of means - Tests based on t and F distributions for mean and variance - Chi-square test- Contingency table (test for independent) - Goodness of fit.

List of Exercise/Experiments using R Programming:

1. Testing of hypothesis for given data using Z - test.
2. Testing of hypothesis for given data using t - test.

UNIT IV 9+6

One way and Two-way classifications - Completely randomized design - Randomized block design -

Latin square design.

List of Exercise/Experiments R Programming:
1. Perform one-way ANOVA test for the given data.
2. Perform two-way ANOVA test for the given data.

UNIT V STATISTICAL QUALITY CONTROL 9+6

Control charts for measurements ($\bar{X}$ and R charts) - Control charts for attributes (p, c and npcharts) - Tolerance limits.

List of Exercise/Experiments using R Programming:
1. Interpret the results for $\bar{X}$-Chart for variable data.
2. Interpret the results for R-Chart for variable data.

OUTCOMES:

Upon completion of the course, the students will be able to:

CO1: Calculate the statistical measures of standard distributions.

CO2: Compute the correlation & regression for two dimensional random variables.

CO3: Apply the concept of testing the hypothesis.

CO4: Implement the concept of analysis of variance for various experimental designs.

CO5: Demonstrate the control charts for variables and attributes.

TEXTBOOKS:

1. R.A. Johnson, I. Miller and J. Freund, "Miller and Freund's Probability and Statistics for Engineers", Pearson Education, Asia, 8th Edition, 2015.
2. J.S. Milton and J.C. Arnold, "Introduction to Probability and Statistics", Tata McGrawHill, 4th Edition, 2017.

REFERENCES:

1. J.L. Devore, "Probability and Statistics for Engineering and the Sciences", Cengage Learning, New Delhi, 9th Edition, 2016.
2. S.M. Ross, "Introduction to Probability and Statistics for Engineers and Scientists", 6th Edition, Elsevier, 2020.
3. M.R. Spiegel, J. Schiller and R.A. Srinivasan, "Schaum‘s Outline of Theory and Problems of Probability and Statistics", Tata McGraw Hill Edition, 2004.
4. R.E.Walpole, R.H.Myers, S.L. Myers and K.Ye, "Probability and Statistics for Engineers and Scientists”. Pearson Education, Asia, 9th Edition, 2012.

R Programming user's Manual

A Brief on R

R is high-level and open-source programming language focused mainly in statistical processing. It is based on the recognized S language and allows the integration with others as C,C++, FORTRAN, Java, Python, etc.

There are many characteristics which have placed it in the elite of the statistical computing software. It is an easy-to-use, flexible, and powerful software with an excellent performance regarding its competitors. Besides it is multiplatform, that is, runs over UNIX, Windows and MacOS. Moreover and last but not least, it is absolutely free; contrasting with the high cost of similar proprietary software.

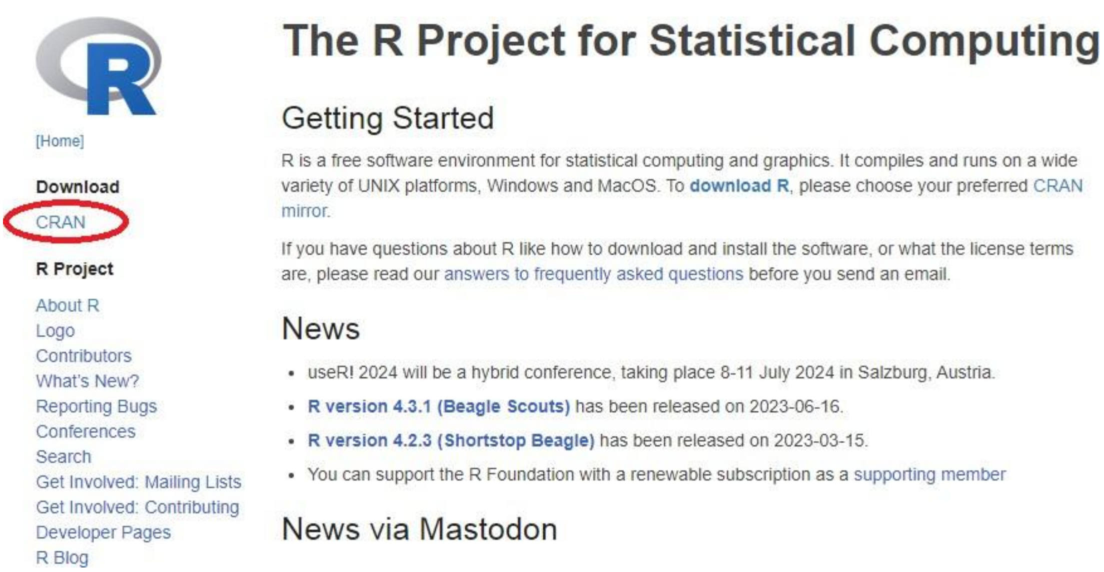
The R software consists barely in a few megabytes including the basic function, which is frequently updated. This philosophy allows a lightly main program kept by the user with only additional applications called packages.These packages are available through the Comprehensive R Archive Network(CRAN).

Applications in R cover a wide range of disciplines such as Bioinformatics,Econometrics,Environmetrics,etc.

A remarkable feature of R is the huge community of users worldwide which have developed an extensive documentation and help sources including a mailing list with keen users.

InstallingR

Open your browser and **visit** www.r-project.org. To **download R**, please click CRAN.

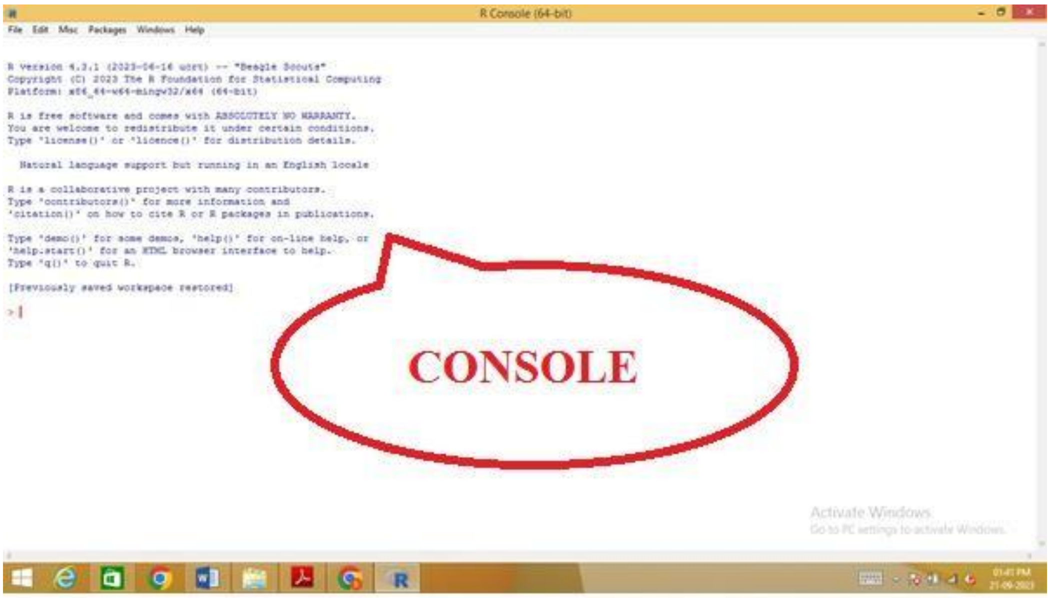


The Comprehensive R Archive Network(CRAN) is available at the following URLs, please choose allocation close to you.

Germany	
https://mirror.howtolearnalanguage.info/	dogado GmbH
https://ftp.fau.de/cran/	Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
https://mirror.dogado.de/cran/	dogado GmbH
https://ftp.gwdg.de/pub/misc/cran/	GWDG Göttingen
https://cran.uni-muenster.de/	University of Münster, Germany
https://mirror.clientvps.com/CRAN/	ClientVPS
https://packages.othr.de/cran/	OTH Regensburg
Greece	
https://ftp.cc.uoc.gr/mirrors/CRAN/	University of Crete
Iceland	
https://cran.hafro.is/	Marine Research Institute
India	
https://cran.icts.res.in/	International Centre for Theoretical Sciences
https://mirror.niser.ac.in/cran/	National Institute of Science Education and Research (NISER)
Indonesia	
https://cran.usk.ac.id/	Universitas Syiah Kuala
Iran	
https://cran.um.ac.ir/	Ferdowsi University of Mashhad
Italy	
https://cran.mirror.garr.it/CRAN/	Garr Mirror, Milano
https://cran.stat.unipd.it/	University of Padua
Japan	
https://cran.ism.ac.jp/	The Institute of Statistical Mathematics, Tokyo
https://ftp.yz.yamagata-u.ac.jp/pub/cran/	Yamagata University

The R installation is very simple.Just download the suitable version for your platform from a desired CRAN at https://cran.icts.res.in/ and install it. When R is opened, appears the R console with a message indicating the following information: the version, the platform, and the important statement that R comes without any warranty, the way to cite R and the packages in publications, etc. Besides that, in contributors () the R-core Team and contributors appear.

In this console the cursor is placed after the > symbol called prompt that indicates availability. On the other hand when + symbol appears, it means that the computation is not completed. Using the Up arrow key it is possible to invoke the last computation.



Add-onPackages

Install BSDA for Unit I by using the command

install.packages("BSDA")

Install dplyr for ANOVA in Unit II using the command

install.packages("dplyr")

Install qcc package for SQC charts for Unit III

install.packages("qcc")

Install pracma package for uni IV

install.packages("pracma")

Note:

Everytime before applying the commands of respective program call the library Function by using the command such as library(BSDA), library(dplyr), library(qcc)etc.

ProgrammingusingR

SimpleManipulation	Rcommand	Output
Addition of numbers	`> 2+3`	`[1] 5`
Assignment operator	`> val<-2`	`> val` `[1] 2`
Combine values into a vector or list	`> p<-c(1.2,3.5,6)`	`> p` `[1] 1.2 3.5 6.0`

All rows and all columns in dataframe	`> df<-matrix(c(2,4,7,3,6,9),`
	`+ nrow=2,ncol=3)`
	`> df[,]`
		`     [,1] [,2] [,3]`
		`[1,]    2    7    6`
		`[2,]    4    3    9`

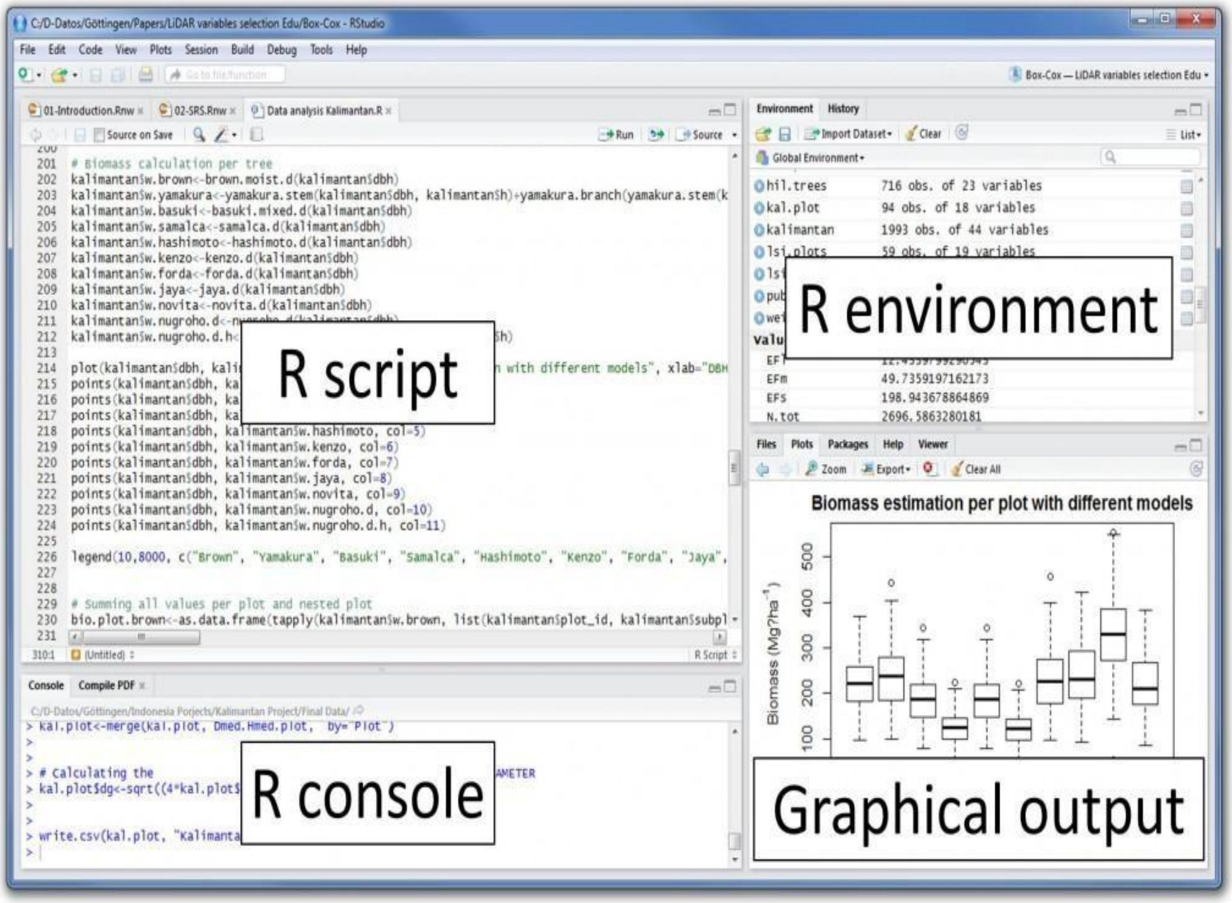
First row and all columns in dataframe	`> df<-matrix(c(2,4,7,3,6,9),`
	`+ nrow=2,ncol=3)`
	`> df[1,]`
		`[1] 2 7 6`

All rows and first and third all columns in dataframe	`> df<-matrix(c(2,4,7,3,6,9),`
	`+ nrow=2,ncol=3)`
	`> df[,c(1,3)]`
		`     [,1] [,2]`
		`[1,]    2    6`
		`[2,]    4    9`

Exp. No. 1	Introduction to R	Date:

AIM: To install R language and present the different types of windows in R language (R Studio).

Out put



Exp.No.2	Finding mean, variance of Binomial distribution	Date:

Problem1:A coin is tossed five times. What is the probability of getting exactly 3 times head? What is the probability of getting atmost 3 times head? Find the cumulative function and find the mean, variance, and standard deviation. Also plot the probability mass function and cumulative function.

Programming command:

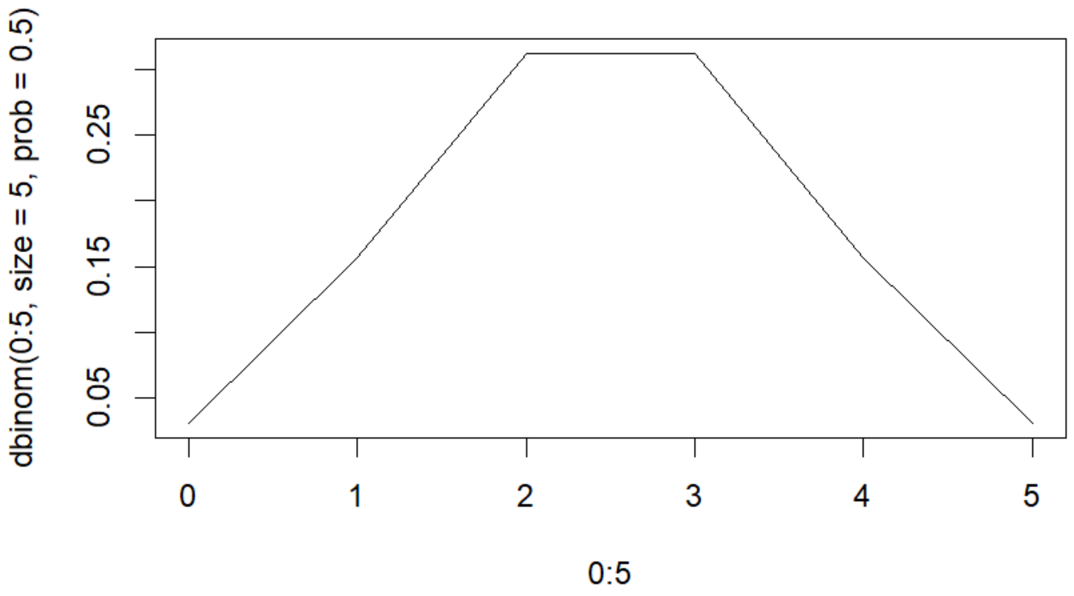
```
> # Number of heads you want to get (exactly 3 times)
> k <- 3
> n<-5
> p<-0.5
> # Calculate the probability using the binomial distribution
PMF
> probability_3_heads <- dbinom(k, size = n, prob = p)
> # Calculate the cumulative probability.
> probability_atmost3_heads <- pbinom(k, size = n, prob = p)
> cumulative_probability <- pbinom(0:5, size = n, prob = p)
```

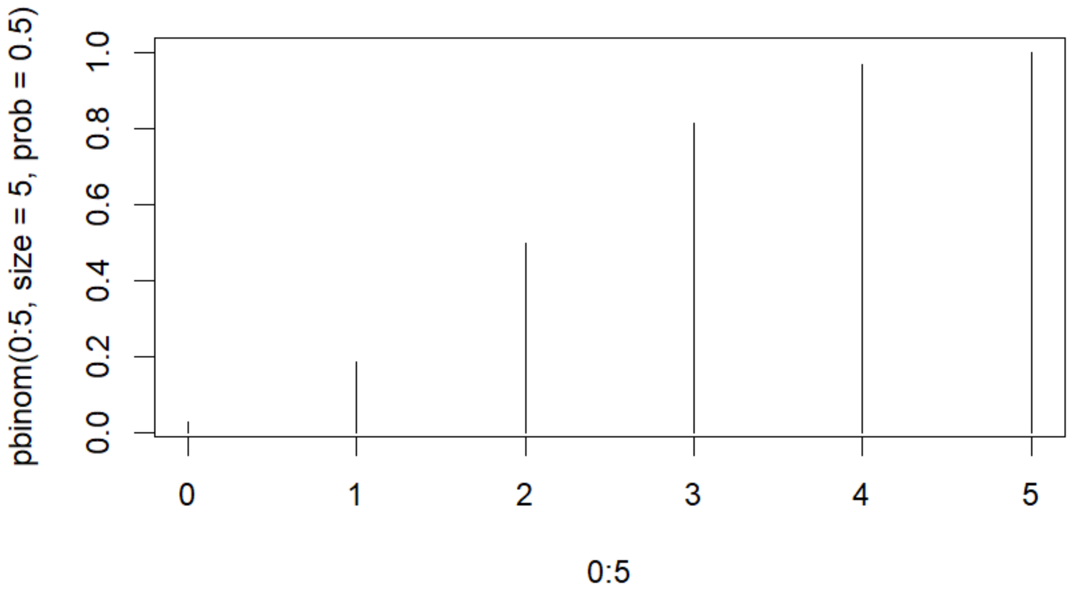
#plot probability mass function and cumulative function

```
> plot(0:5, dbinom(0:5, size = 5, prob = 0.5), type = "l")
> plot(0:5, pbinom(0:5, size = 5, prob = 0.5), type = "h")
>
> # Calculate mean, variance, and standard deviation
> mean_value <- n * p
> variance_value <- n * p * (1 - p)
> std_deviation <- sqrt(variance_value)
> # Display the results
>probability_3_heads
>probability_atmost3_heads
>cumulative_probability
>mean_value
>variance_value
>std_deviation
```

Out put

```
# Display the results
>       probability_3_heads
[1] 0.3125
>       probability_atmost3_heads
[1] 0.8125
>       cumulative_probability
[1] 0.03125 0.18750 0.50000 0.81250 0.96875 1.00000
>       mean_value
[1] 2.5
>       variance_value
[1] 1.25
>     std_deviation
[1] 1.118034
```


dbinom(0:5, size = 5, prob = 0.5)
0.05
0.15
0.25
0
1
2
3
4
5
0:5


pbinom(0:5, size = 5, prob = 0.5)
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
0:5

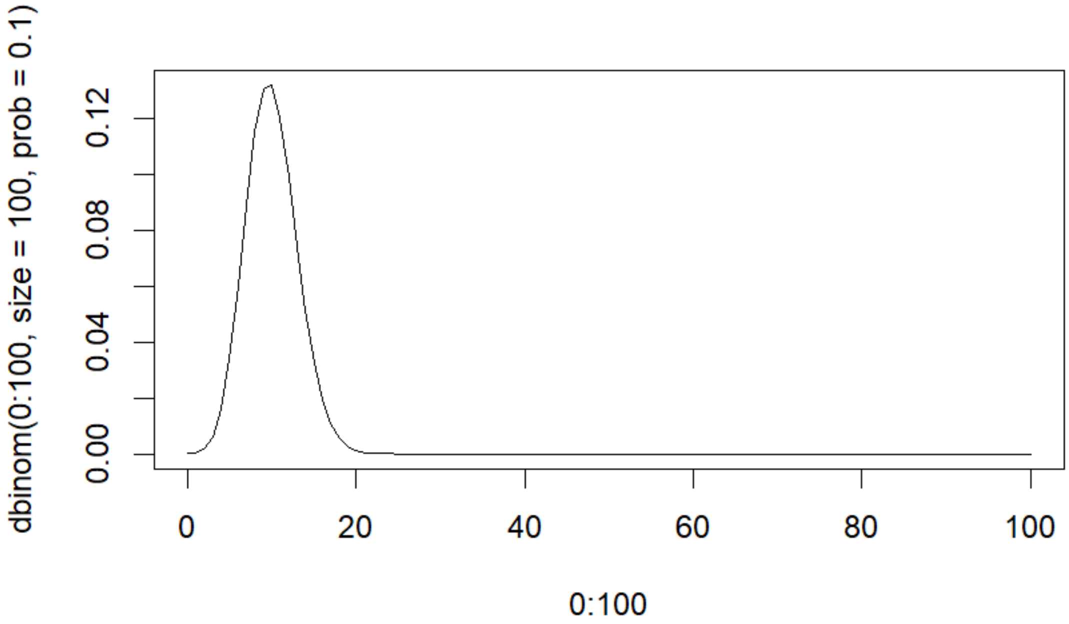
Problem2: If the probability of defective bolts is 0.1, What is the probability that exactly 6 bolts are defective? What is the probability that atmost 6 bolts are defective? Find the cumulative function and find the mean, variance, and standard deviation of the defective bolts in a total of 100 bolts. Also plot the probability mass function and cumulative function.

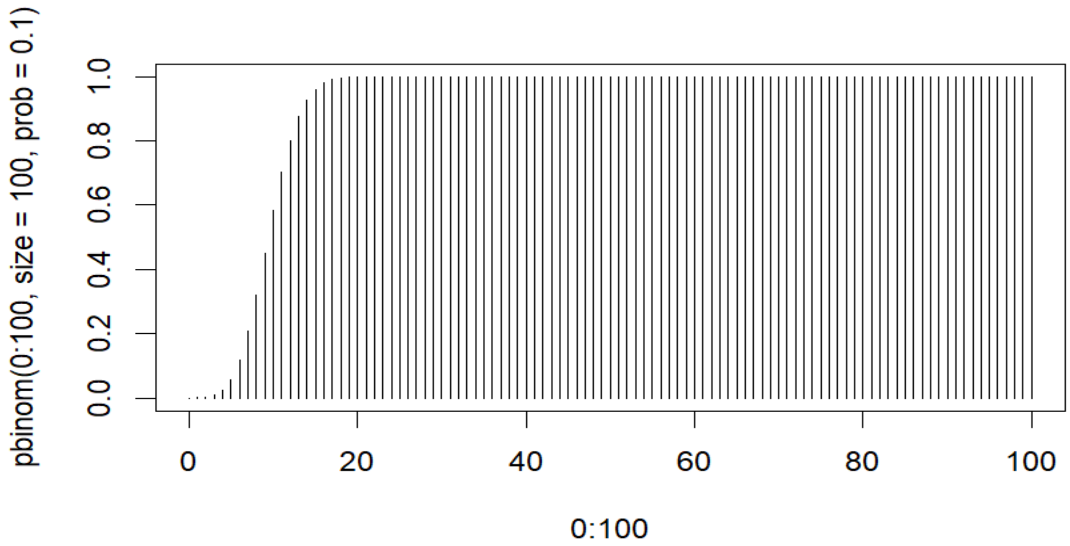
Programming command:

```
> k <- 6
> n<-100
> p<-0.1
> # Calculate the probability using the binomial distribution PMF
>  probability_6_bolts <- dbinom(k, size = n, prob = p)
>  # Calculate the cumulative probability.
>  probability_atmost6_bolts <- pbinom(k, size = n, prob = p)
> cumulative_probability <- pbinom(0:100, size = n, prob = p)
> #plot probability mass function and cumulative function
>  plot(0:100, dbinom(0:100, size = 100, prob = 0.1), type = "l")
>  plot(0:100, pbinom(0:100, size = 100, prob = 0.1), type = "h")
>
> # Calculate mean, variance, and standard deviation
>  mean_value <- n * p
>  variance_value <- n * p * (1 - p)
>  std_deviation <- sqrt(variance_value)
> # Display the results
>probability_6_bolts
>probability_atmost6_bolts
>cumulative_probability
>mean_value
>variance_value
>std_deviation
```

Out put

```
> # Display the results
> probability_6_bolts
[1] 0.05957873
> probability_atmost6_bolts
[1] 0.1171556
> cumulative_probability
  [1] 0.0000265614 0.0003216881 0.0019448847 0.0078364871 0.0237110827
  [6] 0.0575768865 0.1171556154 0.2060508618 0.3208738884 0.4512901654
 [11] 0.5831555123 0.7030331003 0.8018211126 0.8761232074 0.9274270347
 [16] 0.9601094729 0.9794011899 0.9899927207 0.9954192458 0.9980214391
 [21] 0.9991924261 0.9996880820 0.9998858437 0.9999603626 0.9999869272
 [26] 0.9999959001 0.9999987761 0.9999996519 0.9999999056 0.9999999755
 [31] 0.9999999940 0.9999999986 0.9999999997 0.9999999999 1.0000000000
 [36] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [41] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [46] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [51] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [56] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [61] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [66] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [71] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [76] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [81] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [86] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [91] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
 [96] 1.0000000000 1.0000000000 1.0000000000 1.0000000000 1.0000000000
[101] 1.0000000000
> mean_value
[1] 10
> variance_value
[1] 9
> std_deviation
[1] 3
```





Exercise 1: If the probability of defective blub is 0.2, What is the probability that exactly 2 bulbs are defective? What is the probability that atmost 2 bulbs are defective? Find the cumulative function and find the mean, variance, and standard deviation of the defective bulbs in a total of 20 bulbs. Also plot the probability mass function and cumulative function.

Exp.No.3	Finding mean, variance of Normal distribution	Date:

Problem 1: An electric firm manufactures light bulbs that have a life before burns out is normally distributed with mean 800 hours and a S.D of 40 hours. Find the probability that (i) the bulb burns less than 778 hours. (ii)the bulb burns more than 834 hours. (ii) that the bulb burns between 778 and 834 hours.

Programming command:

```
> # calculate the probability less than 778 hours
> Prob_less_778<-pnorm(q=778, mean=800, sd=40, "lower.tail"=T)
> # calculate the probability more than 834 hours
> Prob_more_834<-pnorm(q=834, mean=800, sd=40, "lower.tail"=F)
> #calculate probability between 778 and 834 hours
> prob_778_834 <- pnorm(834, mean = 800, sd = 40) - pnorm(778, mean =
800,   sd = 40)
> #disply the result
> Prob_less_778
> Prob_more_834
> prob_778_834
```

Output

```
#disply the result
>       Prob_less_778
[1] 0.2911597
>       Prob_more_834
[1] 0.1976625
>       prob_778_834
[1] 0.5111778
```

Problem 2: If X is a normal variate with mean 30 and standard deviation 5, find
(a) $P(X \leq 35)$
(b) $P(X \geq 45)$.
(c) $P(26 \leq X \leq 40)$

Programming command:

```
> # calculate the probability less than 35 hours
> Prob_less_35<-pnorm(q=35, mean=30, sd=5, "lower.tail"=T)
> # calculate the probability more than 45 hours
> Prob_more_45<-pnorm(q=45, mean=30, sd=5, "lower.tail"=F)
> #calculate probability between 26 and 40 hours
> prob_26_40 <- pnorm(40, mean = 30, sd = 5) - pnorm(26, mean = 30,sd = 5)
> #disply the result
> Prob_less_35
> Prob_more_45
> prob_26_40
```

Out put

```
#disply the result
>         Prob_less_35
[1] 0.8413447
>         Prob_more_45
[1] 0.001349898
>         prob_26_40
[1] 0.7653945
```

Exercise 1: If X is a normal variate with mean 8 and standard deviation 4, find (a) $P(10 \leq X \leq 15)$ (b) $P(X \geq 15)$, (c) $P(5 \leq X \leq 10)$, c) $P(X \leq 5)$

Exp.No.4	**Finding marginal density functions for discrete random variables.**	**Date:**

Problem1: Let X and Y be two discrete random variable with joint probability mass function $P[X=x, Y=y] = \begin{cases} \frac{x+2y}{21}, & x=1,2,3;\ y=1,2 \\ 0 & ,otherwise \end{cases}$. Find the marginal probabilities of Y=1 and Y=2. Also find conditional probability P(X=2 | Y=1) and P(X=2 | Y=2)

Programming command:

```
> # Define the function
> f <- function(x, y) {
+   return((x + y) / 21)
+ }
> # Given values of x and y
> x_values <- c(1, 2, 3)
> y_values <- c(1, 2)
> # Calculate  probability distribution of f(x, y)
> probability_distribution <- outer(x_values, y_values, FUN = Vectorize(f)
)
> # Calculate joint probability P(X=2, Y=1)
> joint_prob_X2_Y1 <- f(2, 1)
> # Calculate marginal probability P(Y=1), P(Y=2)
> marginal_prob_Y1 <- sum(f(x_values, 1))
> marginal_prob_Y2 <- sum(f(x_values, 2))
> # Calculate conditional probability P(X=2 | Y=1)
> conditional_prob_X2_given_Y1 <- joint_prob_X2_Y1 / marginal_prob_Y1
> conditional_prob_X2_given_Y2 <- joint_prob_X2_Y1 / marginal_prob_Y2
> # Display the result
> probability_distribution
> marginal_prob_Y1
> marginal_prob_Y2
> conditional_prob_X2_given_Y1
> conditional_prob_X2_given_Y2
```

Output

```
> # Display the result
> probability_distribution
          [,1]      [,2]
[1,] 0.0952381 0.1428571
[2,] 0.1428571 0.1904762
[3,] 0.1904762 0.2380952
> marginal_prob_Y1
[1] 0.4285714
> marginal_prob_Y2
[1] 0.5714286
> conditional_prob_X2_given_Y1
[1] 0.3333333
> conditional_prob_X2_given_Y2
[1] 0.25
```

Exercise 1: Let X and Y be two discrete random variable with joint probability mass function $P[X = x, Y = y] = \begin{cases} \frac{x+y}{21}, x = 1,2,3; y = 1,2 \\ 0, otherwise \end{cases}$.Find the marginal probabilities of Y=1 and Y=2. Also find conditional probability P(X=2 | Y=1) and P(X=2 | Y=2).

Exp.No.5	Find the correlation coefficient and Regression lines	**Date:**

Problem 1:

Calculate the Correlation Coefficient and Regression lines for the following

X	**1**	**3**	**5**	**7**	**8**	**10**
Y	**8**	**12**	**15**	**17**	**18**	**20**

Programming command:

```
> library(car)
> x <- c(1,3,5,7,8,10)
> y<- c(8,12,15,17,18,20)
> correlation_coefficient <- cor(x, y)
> cat("Correlation Coefficient (x and y):", correlation_coefficient, "\n")
> regression_x_on_y <- lm(x ~ y)
> coefficients_x_on_y <- coef(regression_x_on_y)
> regression_y_on_x <- lm(y ~ x)
> coefficients_y_on_x <- coef(regression_y_on_x)
> cat("Regression Equation (x on y): y = ", round(coefficients_x_on_y[2], 2), "* x + ",
round(coefficients_x_on_y[1], 2), "\n")
> cat("Regression Equation (y on x): y = ", round(coefficients_y_on_x[2], 2), "* x + ",
round(coefficients_y_on_x[1], 2), "\n")
```

Output:

```
Correlation Coefficient (x and y): 0.9878783
Regression Equation (x on y): y =  0.75 * x +  -5.58
Regression Equation (y on x): y =  1.3 * x +  7.63
```

Problem 2:

Calculate the Correlation Coefficient and Regression lines for the following

X	**65**	**66**	**67**	**67**	**68**	**69**	**70**	**72**
Y	**67**	**68**	**65**	**68**	**72**	**72**	**69**	**71**

Programming command:

```
>library(car)
> x <- c(65,66,67,67,68,69,70,72)
>  y <- c(67,68,65,68,72,72,69,71)
> correlation_coefficient <- cor(x, y)
> cat("Correlation Coefficient (x and y):", correlation_coefficient, "\n")
> regression_x_on_y <- lm(x ~ y)
> coefficients_x_on_y <- coef(regression_x_on_y)
> regression_y_on_x <- lm(y ~ x)
> coefficients_y_on_x <- coef(regression_y_on_x)
> cat("Regression Equation (x on y): y = ", round(coefficients_x_on_y[2], 2), "* x + ",
round(coefficients_x_on_y[1], 2), "\n")
> cat("Regression Equation (y on x): y = ", round(coefficients_y_on_x[2], 2), "* x + ",
round(coefficients_y_on_x[1], 2), "\n")
```

Output:

```
Correlation Coefficient (x and y): 0.6030227
Regression Equation (x on y): y =  0.55 * x +  30.36
Regression Equation (y on x): y =  0.67 * x +  23.67
```

Exercise:

1. Calculate the correlation coefficient and Regression linesfrom the following data:

X	1	3	5	7	8	10
Y	8	12	15	17	18	20

2.Calculate the correlation coefficient and Regression lines from the following data:

X	12	9	8	10	11	13	7
Y	14	8	6	9	11	12	3

3.From the following data, find the co-efficient of correlation between the marks in Mathematics and Statistics and Regression lines

Marks in Mathematics	25	28	35	32	31	36	29	38	34	32
Marks in Statistics	43	46	49	41	36	32	31	30	33	39

Exp.No.6	Testing of hypothesis using Z -test	Date:

Problem1:

Compute the z score to test whether the sample data 88,92,94, 94,96,97,97,97, 99, 99, 83,92,94, 94,96,97,97, 97,99,99,105,109, 109,109,110, 112,94,96,97,97,97,99, 99,83,92,94,94,96,97,97,97,99,99,105,109,109,109,110,112,112,113,114,115
Support the administrator's belief that the mean intelligence test score for all fresh man is greater than109 with sd 10.33.

Programming command:

```
>>library(BSDA)
>> data=c(88, 92, 94, 94, 96, 97, 97, 97, 99,99,
+83, 92, 94, 94, 96, 97, 97, 97, 99, 99, 105,
+109, 109, 109, 110, 112,94, 96, 97, 97, 97,
 +99,99, 83, 92, 94, 94, 96, 97, 97, 97, 99, 99,
 +105, 109, 109, 109, 110,112, 112, 113, 114, 115 )
>>z.test(data, mu=109,alternative="greater", sigma.x = 10.33)
```

Output:

```
            One-sample z-Test

data:  x
z = 4.3401, p-value = 1.424e-05
alternative hypothesis: true mean is not equal to 109
95 percent confidence interval:
 113.0978 119.8466
sample estimates:
mean of x
 116.4722
```

Problem2:

Compute the z score to test whether the sample data

101,103, 112,102, 98,97,93, 105,100,97,107, 93, 94,97, 97,100,110,106, 110, 103,99,93, 98,106,100, 112,105,100,114,97, 110,102, 98,112,99

Can be drawn from a normal population with mean 102and SD 6.

Programming command:

```
>> library(BSDA)

>> data=c(101, 103, 112, 102, 98, 97, 93, 105, 100, 97, 107, 93, 94, 97, 97,
100, 110, 106, 110, 103, 99, 93, 98, 106, 100, 112, 105, 100, 114, 97, 110,
102, 98, 112, 99)

>> z.test(data, mu=102,alternative="two.sided", sigma.x = 6)
```

Output:

```
        One-sample z-Test

data:  x
z = 0, p-value = 1
alternative hypothesis: true mean is not equal to 102
95 percent confidence interval:
 100.0122 103.9878
sample estimates:
mean of x
      102
```

Exercise:

1. Compute the z score to test whether the sample data

85,85,78,78, 92,94,91,85, 72,97,84,95,99, 80,90,88,95, 90,96,89,84, 88,88,90,92,

93, 91, 85, 80, 93,97, 100,93, 91, 90, 87, 94, 83,92, 95

That the mean intelligence test score for all freshman is greater than100 with sd 6.19.

2. Compute the z score to test whether the sample data

5,5, 8,7, 2,9, 1,8,7, 9,8,5,9, 8,0,8, 9,9,6, 8, 4,8, 8,9,2, 3,1,5,8,3, 7, 2,3,1, 4,7,4,

3,2,5 that the mean intelligence test score for all freshman is greater than 5with sd1.828.

Exp.No.7	Testing of hypothesis using t-test	Date:

Problem1:

A random sample of 10 boys had the followingI.Q's:70,120,110,101,88,83,95,98,107,100. Do these data support the assumption of a population mean I.Q of 100? Find a reasonable range in which most of the mean I.Q values of samples of 10 boys lie.

Programming command:

```
library(BSDA)

>> IQ=c (70, 120, 110, 101, 88, 83, 95, 98, 107, 100)

>> t.test(IQ, mu=100,alternative=" two.sided",conf.level=0.95)
```

Output:

```
        One Sample t-test

data:  x
t = -0.62034, df = 9, p-value = 0.5504
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
  86.98934 107.41066
sample estimates:
mean of x
     97.2
```

Problem2:

The marks obtained by a group of 9 regular course students and another group of 11part-time course students in a test are given below.

Regular	56	62	63	54	60	51	67	69	58		
Part-time	62	70	71	62	60	56	75	64	72	68	66

Perform a hypothesis test whether the marks obtained by regular students and part-time students differ significantly at 5%.

Programming command:

```
> regular<-c(56,62,63,54,60,51,67,69,58)
> part.time<-c(62,70,71,62,60,56,75,64,72,68,66)
> t.test(x=regular,y=part.time,alternative="two.sided",conf.level=0.95,var.equal=TRUE)
```

Output:

```
        Two Sample t-test

data:  regular and part.time
t = -2.2856, df = 18, p-value = 0.03462
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -11.5151239  -0.4848761
sample estimates:
mean of x mean of y 
       60        66 
```

Exercise:

1. Sandal powder is packed into packets by a machine. A random sample of 12 packets is drawn and their weights are found to be (inkg) 0.49,0.48,0.47,0.48,0.49,0.50,0.51,0.49,0.50,0.51and 0.48.Test if the average weight of the packing can be taken as 0.5kg.

2. Two independent samples of sizes 8 and 7 contained the following values

Sample I	19	17	15	21	16	18	16	14
Sample II	15	14	15	19	15	18	16	

Is the difference between the sample means significant?

Exp.No.8	Perform one way ANOVA test	Date:

Problem1:

Compute the one way analysis of variance table for the following data

A	B	C	D
20	25	24	23
19	23	20	20
21	21	22	20

Programming command:

```
> table<-data.frame(b1=c(20,19,21),
+ b2=c(25,23,21),
+ b3=c(24,20,22),
+ b4=c(23,20,20))
> table_stacked<-stack(table[,1:4])
> anova<-aov(values~ind,data=table_stacked)
> summary(anova)
```

Output:

```
            Df Sum Sq Mean Sq F value Pr(>F)
ind          3     15       5   1.667   0.25
Residuals    8     24       3
>
```

Problem2:

Compute the analysis of variance table to test the difference among the sample means at three positions are significant.

Position 1	90	82	79	98	83	91	
Position 2	105	89	93	104	89	95	86
Position 3	83	89	80	94			

Programming command:

```
> table<-data.frame(p1=c(90,82,79,98,83,91,NA),
+ p2=c(105,89,93,104,89,95,86),
+ p3=c(83,89,80,94,NA,NA,NA))
> table_stacked<-stack(table[,1:3])
> anova<-aov(values~ind,data=table_stacked)
> summary(anova)
```

Output:

```
            Df Sum Sq Mean Sq F value Pr(>F)
ind          2  234.5  117.23   2.333  0.134
Residuals   14  703.5   50.25
4 observations deleted due to missingness
```

Exercise:

1. Compute the one way analysis of variance table for the following data

A	B	C	D
8	6	14	20
9	8	12	22
11	10	18	25
12	4	9	23

2. Compute the analysis of variance table to test the homogeneity of the mean lives of the four brands of electric lamps.

A	1610	1610	1650	1680	1700	1720	1800	
B	1580	1640	1700	1750				
C	1460	1550	1600	1620	1640	1660	1740	1820
D	1510	1520	1530	1570	1600	1680		

Exp.No.9	Perform two way ANOVA test	Date:

Problem1:

Compute the two-way analysis of variance table for the following data

Salesmen

Season		A	B	C	D
	Summer	45	40	38	37
	Winter	43	41	45	38
	Monsoon	39	39	41	41

Programming command:

```
> sales<-c(45,40,38,37,43,41,45,38,39,39,41,41)
> season<-c("A","A","A","A","B","B","B","B","C","C","C","C")
> class(season)
[1] "character"
> season<-factor(season,levels=c("A","B","C","D"))
> salesmen<-c("1","2","3","4","1","2","3","4","1","2","3","4")
> salesmen<-factor(salesmen,levels=c("1","2","3","4"))
> anova2way<-aov(sales~season+salesmen)
> summary(anova2way)
```

Output:

```
            Df  Sum Sq  Mean Sq  F value  Pr(>F)
season       2    8.17    4.083    0.535   0.611
salesmen     3   22.92    7.639    1.000   0.455
Residuals    6   45.83    7.639
```

Problem2:

Compute the two way analysis of variance table for the following data

Machine type

Workers		A	B	C	D
	W1	44	38	47	36
	W2	46	40	52	43
	W3	34	36	44	32
	W4	43	38	46	33
	W5	38	42	49	39

Programmingcommand:

```
> treat<-c(44,38,47,36,46,40,52,43,34,36,44,32,43,38,46,33,38,42,49,39)
> block<-c("B1","B1","B1","B1","B2","B2","B2","B2","B3","B3","B3","B3","B4","B4$
> class(block)
[1] "character"
> block<-factor(block,levels=c("B1","B2","B3","B4","B5"))
> treatment<-c("1","2","3","4","1","2","3","4","1","2","3","4","1","2","3","4",$
> treatment<-factor(treatment,levels=c("1","2","3","4"))
> anova2way<-aov(treat~block+treatment)
> summary(anova2way)
```

Output:

```
            Df Sum Sq Mean Sq F value   Pr(>F)
block        4  161.5   40.38   6.574  0.00485 **
treatment    3  338.8  112.93  18.388 8.78e-05 ***
Residuals   12   73.7    6.14
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Exercise:

1. Compute the two-way analysis of variance table for the following data

Treatment

		A	B	C	D
Plots of land	I	38	40	41	39
	II	45	42	49	36
	III	40	38	42	42

2. Compute the two way analysis of variance table for the following data

		Engine1	Engine2	Engine1
Detergents	A	45	43	51
	B	47	46	52
	C	48	50	55
	D	42	37	49

Exp.No.10	Plot $\overline{X}$ chart for the variable data	Date:

Problem1:

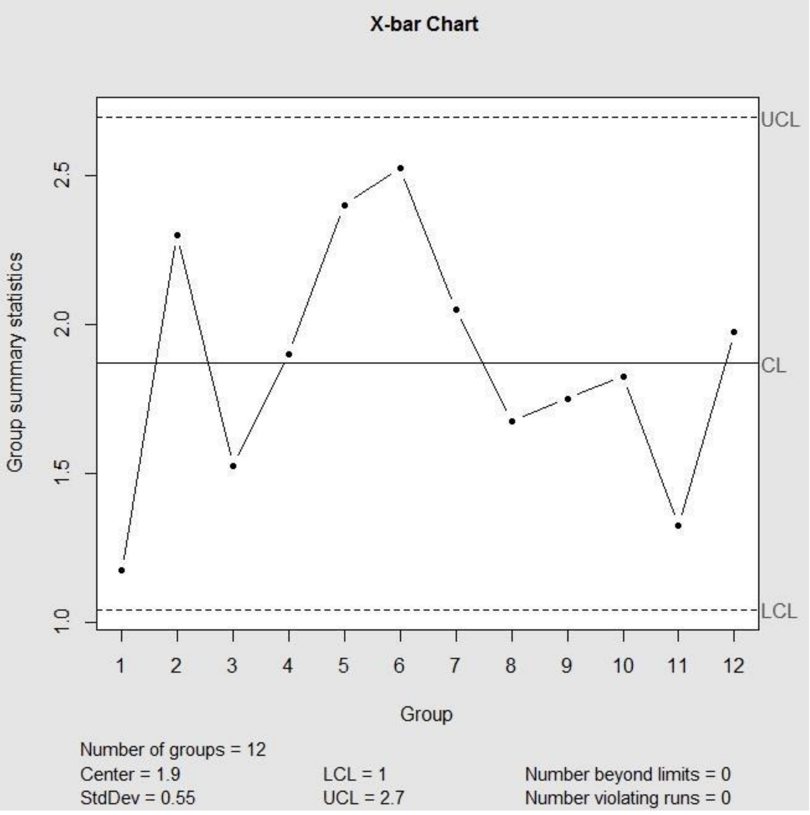
Plot $\overline{X}$ chart for the following 12 samples each having 4 units.

Sample No	1	2	3	4	5	6	7	8	9	10	11	12
Measurements in(mm)	1	1.4	1.6	2.9	1.7	2.6	2.3	1.9	1.7	1.8	0.8	2.0
	1.4	2.3	1.0	2.0	3.6	2.8	2.1	1.6	2.2	2.0	1.5	2.5
	1.3	2.8	1.5	0.5	2.5	3.2	2.1	1.8	1.9	1.5	2.1	1.6
	1	2.7	2.0	2.2	1.8	1.5	1.7	1.4	1.2	2.0	0.9	1.8

Programmingcommand:

```
> library(qcc)
  __ _  ___ ___
 / _` |/ __/ __|   Quality Control Charts and
| (_| | (_| (__    Statistical Process Control
 \__, |\___\___|
    |_|            version 2.7
Type 'citation("qcc")' for citing this R package in publications.
> chartdata2<-read.table(header=FALSE,text="
+ 1 1.4 1.3 1
+ 1.4 2.3 2.8 2.7
+ 1.6 1 1.5 2
+ 2.9 2 0.5 2.2
+ 1.7 3.6 2.5 1.8
+ 2.6 2.8 3.2 1.5
+ 2.3 2.1 2.1 1.7
+ 1.9 1.6 1.8 1.4
+ 1.7 2.2 1.9 1.2
+ 1.8 2 1.5 2
+ 0.8 1.5 2.1 0.9
+ 2 2.5 1.6 1.8")
> xbar_chart2 <- qcc(data = chartdata2, type = "xbar", sizes = 4, title = "X-bar Chart ", digits = 2, plot = TRUE)
```

Output:



Problem2:

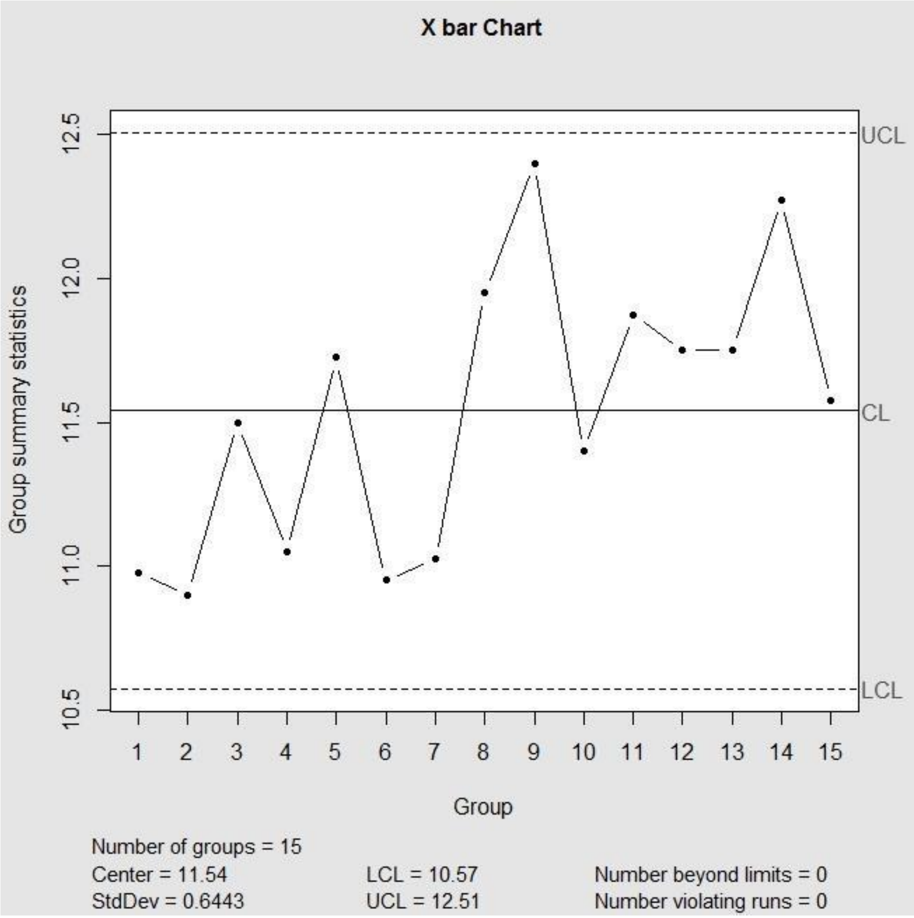
Plot $\overline{X}$ chart for the following weights of 15 samples of 4 boxes drawn randomly.

Sample No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Box1	10.0	10.3	11.5	11.0	11.3	10.7	11.3	12.3	11.0	11.3	12.5	11.9	12.1	11.9	10.6
Box2	10.2	10.9	10.7	11.1	11.6	11.4	11.4	12.1	13.1	12.1	11.9	12.1	11.1	12.1	11.9
Box3	11.3	10.7	11.4	10.7	11.9	10.7	11.1	12.7	13.1	10.7	11.8	11.6	12.1	13.1	11.7
Box4	12.4	11.7	12.4	11.4	12.1	11.0	10.3	10.7	12.4	11.5	11.3	11.4	11.7	12.0	12.1

Programmingcommand:

```
> chartdata<-read.table(header=FALSE,text="
+ 10 10.2 11.3 12.4
+ 10.3 10.9 10.7 11.7
+ 11.5 10.7 11.4 12.4
+ 11 11.1 10.7 11.4
+ 11.3 11.6 11.9 12.1
+ 10.7 11.4 10.7 11
+ 11.3 11.4 11.1 10.3
+ 12.3 12.1 12.7 10.7
+ 11 13.1 13.1 12.4
+ 11.3 12.1 10.7 11.5
+ 12.5 11.9 11.8 11.3
+ 11.9 12.1 11.6 11.4
+ 12.1 11.1 12.1 11.7
+ 11.9 12.1 13.1 12
+ 10.6 11.9 11.7 12.1")
> Xbar_chart <- qcc(data = chartdata, type = "xbar", sizes = 4, title = "X bar Chart ", digits = 4, plot = TRUE)
```

Output:



Exercise:

1. In a production line,measurements were made by taking 15 samples,each containing 4 numbers. Plot $\overline{X}$ chart.

SampleNo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Measurements	1.7	0.8	1	0.4	1.4	1.8	1.6	2.5	2.9	1.1	1.7	4.6	2.6	2.3	1.9
	1.5	1.4	0.6	2.3	2.0	1.0	1.6	1.6	2.0	1.1	3.6	2.8	1.8	2.1	1.6
	1.9	2.1	1.0	0.7	2.8	1.1	1.5	1.8	0.5	3.1	2.5	3.5	3.2	2.1	1.8
	0.9	0.9	1.3	0.2	2.7	0.1	2.0	1.2	2.2	1.6	1.8	1.9	1.5	1.7	1.4

2. Track Bicycle Parts manufactures precision ball bearings for wheel hubs, bottom brackets, headsets, and pedals. The output of the 5-mm bearings used in front wheel hubs for 18 hours sampled 5 bearings are given below.

Hours	1	2	3	4	5	6	7	8	9	10
BearingDiameters(mm)	5.03	4.97	5.02	4.92	5.01	5	4.94	5	4.99	5.03
	5.06	4.94	4.98	4.93	4.99	4.95	4.91	4.98	5.01	4.96
	4.86	5.09	4.94	4.90	4.93	5.10	5.05	5.05	4.93	4.92
	4.90	4.78	4.95	4.92	5.06	4.85	5.07	4.96	5.10	5.01
	4.95	4.88	4.80	4.96	5.01	4.91	4.88	4.97	4.98	4.93
4Hours	11	12	13	14	15	16	17	18		
BearingDiameters(mm)	5.02	5.09	4.9	5.04	5.09	5.10	4.97	5.01		
	4.88	5.01	4.93	4.96	4.90	5.01	5.10	4.99		
	5.00	5.13	4.97	5.15	5.04	5.04	5.12	5.06		
	4.98	4.89	4.98	5.04	5.19	5.05	4.92	5.04		
	5.09	5.02	5.12	5.02	5.03	5.02	5.04	5.12		

Construct an $\overline{X}$ chart to check whether the production of 5-mm bearings is in-control.

Exp.No.11	Plot R chart for the variable data	Date:

Problem1:

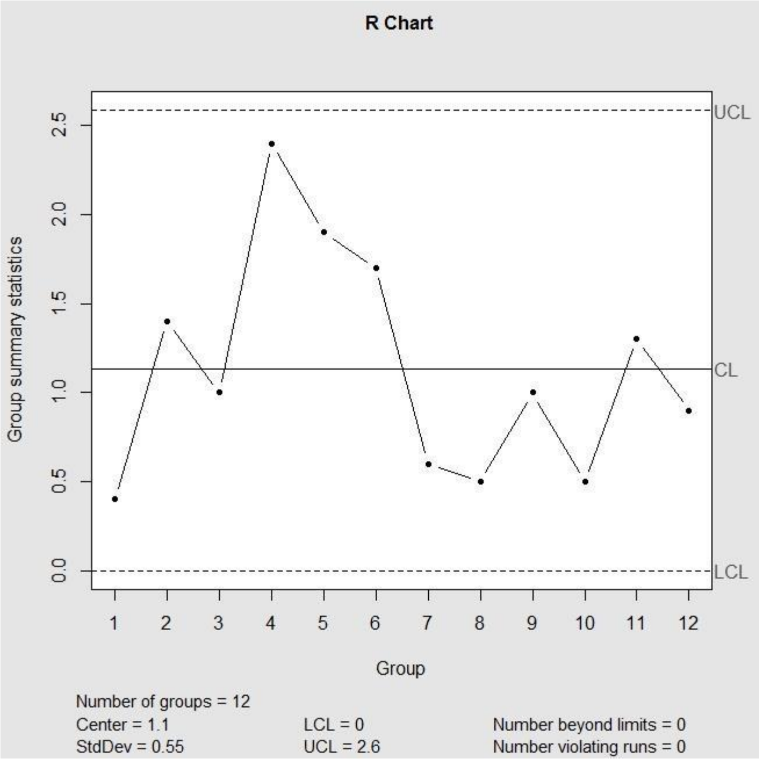
Plot R chart for the following weights of 15 samples of 4 boxes drawn randomly.

Sample No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Box1	10.0	10.3	11.5	11.0	11.3	10.7	11.3	12.3	11.0	11.3	12.5	11.9	12.1	11.9	10.6
Box2	10.2	10.9	10.7	11.1	11.6	11.4	11.4	12.1	13.1	12.1	11.9	12.1	11.1	12.1	11.9
Box3	11.3	10.7	11.4	10.7	11.9	10.7	11.1	12.7	13.1	10.7	11.8	11.6	12.1	13.1	11.7
Box4	12.4	11.7	12.4	11.4	12.1	11.0	10.3	10.7	12.4	11.5	11.3	11.4	11.7	12.0	12.1

Programming command:

```
> chartdata<-read.table(header=FALSE,text="
+ 10 10.2 11.3 12.4
+ 10.3 10.9 10.7 11.7
+ 11.5 10.7 11.4 12.4
+ 11 11.1 10.7 11.4
+ 11.3 11.6 11.9 12.1
+ 10.7 11.4 10.7 11
+ 11.3 11.4 11.1 10.3
+ 12.3 12.1 12.7 10.7
+ 11 13.1 13.1 12.4
+ 11.3 12.1 10.7 11.5
+ 12.5 11.9 11.8 11.3
+ 11.9 12.1 11.6 11.4
+ 12.1 11.1 12.1 11.7
+ 11.9 12.1 13.1 12
+ 10.6 11.9 11.7 12.1")
> R_chart <- qcc(data = chartdata, type = "R", sizes = 4, title = "R Chart ", digits = 4, plot = TRUE)
```

Output:



Problem2:

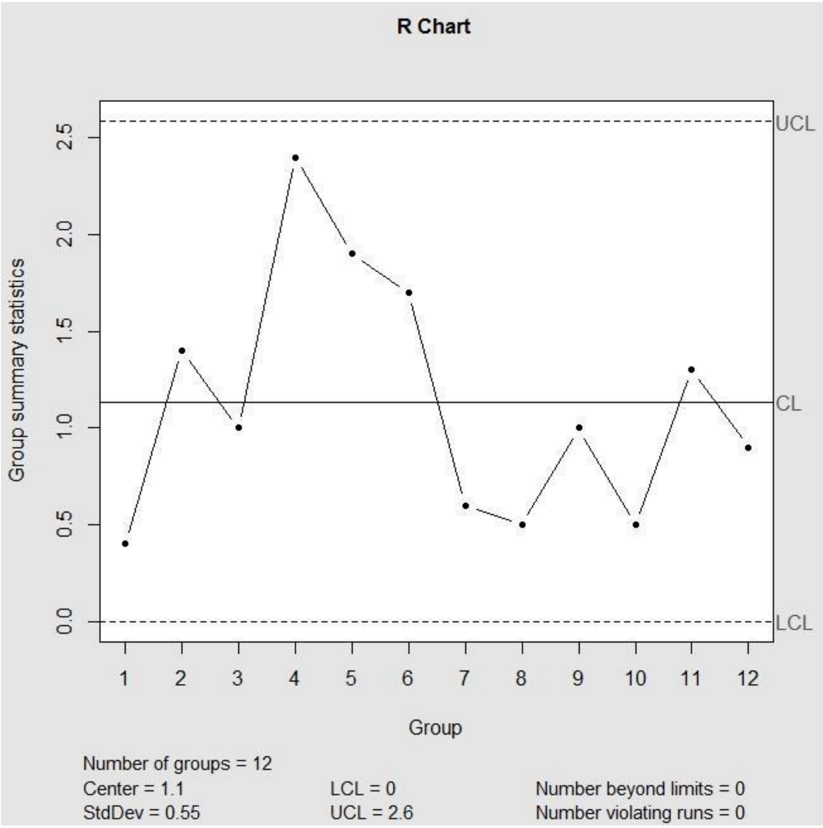
Plot R chart for the following 12 samples each having 4 units.

Sample No	1	2	3	4	5	6	7	8	9	10	11	12
Measurements in(mm)	1	1.4	1.6	2.9	1.7	2.6	2.3	1.9	1.7	1.8	0.8	2.0
	1.4	2.3	1.0	2.0	3.6	2.8	2.1	1.6	2.2	2.0	1.5	2.5
	1.3	2.8	1.5	0.5	2.5	3.2	2.1	1.8	1.9	1.5	2.1	1.6
	1	2.7	2.0	2.2	1.8	1.5	1.7	1.4	1.2	2.0	0.9	1.8

Programming command:

```
> chartdata<-read.table(header=FALSE,text="
+ 1 1.4 1.3 1
+ 1.4 2.3 2.8 2.7
+ 1.6 1 1.5 2
+ 2.9 2 0.5 2.2
+ 1.7 3.6 2.5 1.8
+ 2.6 2.8 3.2 1.5
+ 2.3 2.1 2.1 1.7
+ 1.9 1.6 1.8 1.4
+ 1.7 2.2 1.9 1.2
+ 1.8 2 1.5 2
+ 0.8 1.5 2.1 0.9
+ 2 2.5 1.6 1.8")
> R_chart <- qcc(data = chartdata, type = "R", sizes = 4, title = "R Chart ", digits = 2, plot = TRUE)
>
```

Output:



Exercise:

1. Track Bicycle Parts manufactures precision ball bearings for wheelhubs, bottom brackets, headsets and pedals. The output of the 5-mm bearings used in front wheel hubs for 18 hours sampled 5 bearings are given below.

Hours	1	2	3	4	5	6	7	8	9	10
Bearing Diameters (mm)	5.03	4.97	5.02	4.92	5.01	5	4.94	5	4.99	5.03
	5.06	4.94	4.98	4.93	4.99	4.95	4.91	4.98	5.01	4.96
	4.86	5.09	4.94	4.90	4.93	5.10	5.05	5.05	4.93	4.92
	4.90	4.78	4.95	4.92	5.06	4.85	5.07	4.96	5.10	5.01
	4.95	4.88	4.80	4.96	5.01	4.91	4.88	4.97	4.98	4.93
Hours	11	12	13	14	15	16	17	18		
Bearing Diameters (mm)	5.02	5.09	4.9	5.04	5.09	5.10	4.97	5.01		
	4.88	5.01	4.93	4.96	4.90	5.01	5.10	4.99		
	5.00	5.13	4.97	5.15	5.04	5.04	5.12	5.06		
	4.98	4.89	4.98	5.04	5.19	5.05	4.92	5.04		
	5.09	5.02	5.12	5.02	5.03	5.02	5.04	5.12		

Construct R chart to check whether the production of 5-mm bearings is in-control.

2. In a production line, measurements were made by taking 15 samples, each containing 4 numbers. Plot R chart.

Sample No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Measurements	1.7	0.8	1	0.4	1.4	1.8	1.6	2.5	2.9	1.1	1.7	4.6	2.6	2.3	1.9
	1.5	1.4	0.6	2.3	2.0	1.0	1.6	1.6	2.0	1.1	3.6	2.8	1.8	2.1	1.6
	1.9	2.1	1.0	0.7	2.8	1.1	1.5	1.8	0.5	3.1	2.5	3.5	3.2	2.1	1.8
	0.9	0.9	1.3	0.2	2.7	0.1	2.0	1.2	2.2	1.6	1.8	1.9	1.5	1.7	1.4