Practical - 01

R is a software for statistical analysis and data computing.

It is an effective data handling software and outcome storage is possible.

It is capable of graphical display.

R is a free software.

Q.1] solve the following.

	code:
1. $4+6+8 \div 2-5$	1] 4+6+8/2 [1] 14
2. $2^2+\lvert -3\rvert+\sqrt{45}$	2] 2^2+abs(-3)+sqrt(45) [1] 13.7082
3. $5^3+7\times5\times8+46/5$	3] 5^3+7*5*8+46/5 [1] 414.2
$\sqrt{4^2+5\times3+7/6}$	4] sqrt(4^2+5*3+7/6) [1] 5.671567
Round off $46\div7+9\times8$	5] 46/7+9*8 [1] 78.57143

Q.2] solve the following :

① c(2,3,5,7) * 2
-> [1] 4 6 10 14

② c(2,3,5,7) * c(2,3)
-> [1] 4 9 10 21

③ c(2,3,5,7) * c(2,3,6,2)
-> [1] 4 9 30 14

④ c(1,6,2,3) * c(-2,-3,-4,-1)
-> [1] -2 -18 -8 -3

⑤ c(2,3,5,7) ^ 2
-> [1] 4 9 25 49

⑥ c(4,6,8,9,4,5) ^ c(1,2,3)
-> [1] 4 36 512 9 16 125

⑦ c(6,12,7,5) / c(4,5)

Q.3] Solve the following:

$x = 20$, $y = 30$, $z = 2$

find i) $x^2 + y^3 + z$

ii) $\sqrt{x^2 + y}$

iii) $x^2 + y^2$

) > a = [illegible] + y^3 + z

>y = matric (nrow = 3, ncol = 3, data = c({10, 12, 15,
-5, -4, -6, 7, 9, 5)

> y

	[,1]	[,2]	[,3]
[1,]	10	-5	7
[2,]	12	-4	9
[3,]	15	-6	5

> 2^x + 3^

	[,1]	[,2]	[,3]
[1,]	38	-19	33
[2,]	50	-12	41
[3,]	63	-28	21

> x + y

	[,1]	[,2]	[,3]
[1,]	14	-7	13
[2,]	19	-4	15
[3,]	24	-11	8

Q.6) Marks of statistics ~~of~~ computer science student

59, 20, 35, 24, 46, [illegible] 45, 27, 22, 27,
58, 54, 40, 50, 3[illegible] 35 → 39

Practical-2 :

Aim: Probability distribution:

i) check weather the following are p.m.

1)

x	$P(x)$
0	0.1
1	0.2
2	-0.5
3	0.4
4	0.3
5	0.5

=> since, $P(2) = -0.5$, be pmf because $P(x) \geq 0 \ \forall x$

ii)

x	1	2	3	4	5
$P(x)$	0.2	0.2	0.3	0.2	0.2

```
> prob = c(0.2,0.2,0.3,0.2,0.2)
> sum (prob)
[1] 1.1
```

It cannot be p.m.f in p.m.f, $P(x)=1$

iii)

x	10	20	30	40	50
$P(x)$	0.2	0.2	0.35	0.15	0.1

```
> prob = c (0.2, 0.2, 0.35, 0.15, 0.1)
> sum(prob)
[1] 1
```

It is a p.m.f [illegible] $(x)=1$ and $P(x) \geq 0 \ \forall$

120

Q.2]

x	1	2	3	4	5	6
P(x)	0.15	0.25	0.1	0.20	0.2	0.1

```
> prob= c (0.15, 0.25, 0.1, 0.2, 0.2, 0.
> sum (prob)
> cumsum (prob)
[1] 0.15   0.40   0.50   0.70   0.90   1.0
> x = c (1, 2, 3, 4, 5, 6)
> plot (x, cumsum (prob), "s", xlab=
  "values", ylab " cumsum freq", main
  "CDF ")
```

220

(1) $\int_0^1 f(x)\, dx$

$= \int_0^1 3x^2 \, dx$

$\left[\frac{3x^3}{3} \right]_0^1$

$= \left[x^3 \right]_0^1$

$= 1$

since integration is equal to 1,
It is a p.d.f.

AM

Scanned with CamScanner

solutions

1) > dbinom (10, 100, 0.1)
[1] 0.1318653

2)
1) exactly 4 correct answers
> dbinom (4, 12, 0.2)
[1] 0.1328756

2) Atmost 4 correct answer
> pbinom (4, 12, 0.2)
[1] 0.9274445

3) more than 5 correct answer
> 1 - pbinom (5, 12, 0.2)
[1] 0.01940528

[3] dbinom (0:5, 5, 0.1)
[1] 0.59049 0.32805 0.07290 0.0081
0.00045

[4]
i) P(X=5) = > db ... 0.25) [1] 0.103
ii) P(X≤5) = > pb ... 5) [1] 0.9...
iii) P(X>7) = ... 2, 0.25) [1] 0.8...
iv) P(5<X≤7) ... b ... 25) [1] 0.040

Practical - 3

Topic: Binomial distribution

Commands
i) P(X = x) = dbinom (x, n, p)
ii) P(X ≤ x) = pbinom (x, n, p)
iii) P(X > x) = 1 - pbinom (x, n, p)
iv) If x is unknown and $p_1 = P(X \le x)$
-> qbinom (p_1, n, p)

Q] find the probability of exactly 10 sucess in 100 trials. with p = 0.1

suppose there are 12 MCQ. Each question has 5 options out of which one is correct. find the probability of having
exactly 4 correct answer
Atmost 4 correct answers
more than 5 correct answer.

find the complete distribution ...
n=5, p = 0.1

n=12, p = 0.25 find the prob...
i) P(X=5) ii) P(X ≤ 5) ...
iv) P(5 < X < 7)

Q] Probability of a salesman making a sale to a customer 0.15. Find the probability of
i) no sale out of 10 customer
ii) more than 3 sale out of 10 20 customer.

Q] A salesman has a 20% probability of making a sale out of 30 customer what min no of sales he can make with 88% probability

Q] X follows binomial distribution with $n=10$, $p=0.3$, plot the graph of pmf and cdf

220.

> plot (x, prob, "n")

> plot (x, comprob, "c").

> P3 = 1- pnorm (14, 12, 3)

[1] 0.2524925

> cat (" P (x) (u) = ", P3)

(6)

> rnorm (5, 12, 3)

[1] 10.512180 11.924144 15.11409 10.362

Q.2] x follows normal distribution with u=10, find

(1) P(x ≤ 7) 2) P(5 < x < 12) 3) P(x > 12) 4) generate 10 observation

(5) find k that P(x < k) = 0.4

=>

1) P(x ≤ 7)

P1 = pnorm (7, 10, 2)

> P1

[1] 0.0668072

2) P(5 < x < 12)

> P2 = pnorm (12, 10, 2) - pnorm (5, 10, 2)

> P2

[1] 0.835135

) P(x > 12)

P3 = 1- pnorm (12, 10, 2)

P3

0.158655 3

generate 10 observation

norm (10, 10, 2)

7.324489 8.308746 8.394823 11.990873

10.754449 13.33875 10.266751

020

p(25 < x < 35)
> p3 = pnorm(35,30,10) - pnorm(25,30,10)
> p3
[1] 0.3829 249

4) qnorm(0.6,30,10)
[1] 32.533 47

Q3) solution
n=5 , μ = 15 , σ = 4
rnorm(5,15,4)
> am = mean(x)
> am
[1] 16.31588
> me=median(x)
> me
[1] 15.58 446

> n=5
> variance = (n-1) * var(x)/n
> sd = sqrt (variance)
> sd
[1] 2.33258 3.

> cat ("sample mean = ", am)
> cat ("sample median = ", me)
> cat (" sample s.d =", sd)

Q.3) Generate 5 random no from normal distribution and mean μ=15, sd=4, hence sample mean, mediation, SD and ~~variance~~

Q.4) x follows N(30,100)
μ = 30, σ = 10
find 1) P(x ≤ 40) 2) P(x>35) 3) P(25<x<35) 4) find k P(x<k)

=> 1) P(x ≤ 40)
> p1 = pnorm(40,30,10)
> p1
[1] 0.841344 7
2) P(x>35)
> p2 = 1 - pnorm(35,30,10)
> p2
[1] 0.3085 375.
>

ii) sample mean = 15.58444

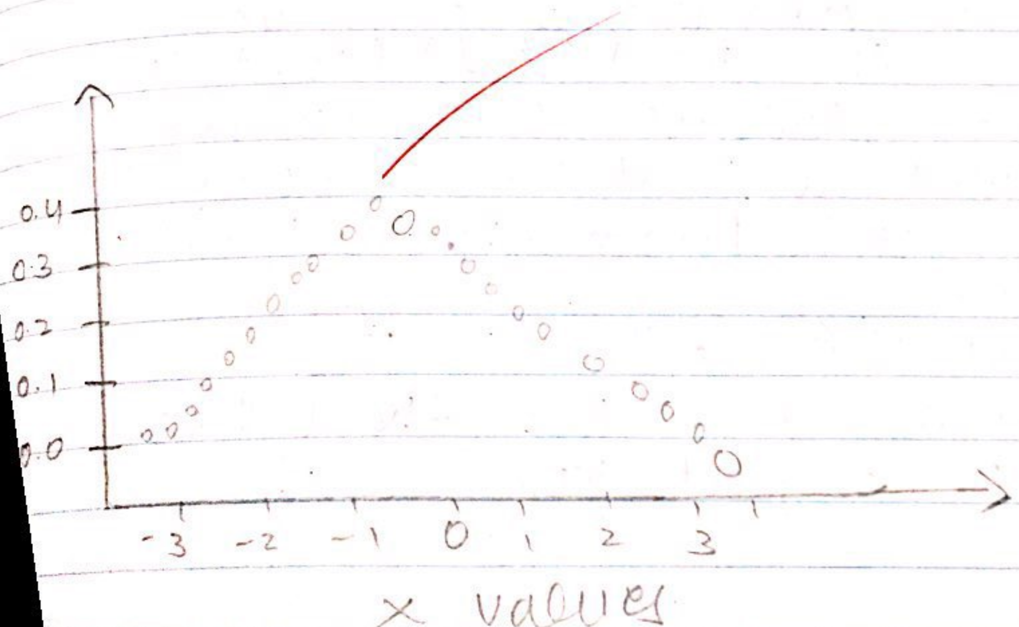
iii) sample variance = 2.3332583

plot the standard normal graph

x = seq(-3, 3, by = 0.1)

y = dnorm(x)

plot(x, y, xlab = "x values", ylab = "Probability", main = "standard normal graph")



Practical - 05

Aim: Normal and T. test

Q1] Test the Hypothesis

$H_0 : \mu = 15$

$H_1 : \mu \neq 15$

Random sample of 400 size of drawn and it is calculated. The sample mean is 14 and standard deviation is 3. Hypothesis at 5% level of significance.

```
-> > mo = 15 , mx = 14 , sd = 3 , n = 400
> zcal = (mx - mo) / (sd) (sqrt (n)))
> cat (" calculated value at z is =", zcal
calculated value at z is -6.66667
> pvalue = 2 * (1 - pnorm (abs (zcal)))
> pvalue
[1] 2.616796 e-11
```

∴ value of P is less than 0.05.

Test the Hypothesis

Test $\mu = 10$ $H_0 : \mu = 10$

$\mu \neq 10$ $H_1 : \mu \neq 10$

A random sample of size 400 is drawn with sample mean 10.2 and s.d = 2.5.

Test the sample mean 10.2 and s.d = 2.5

Test the hypothesis at 5% level of significance.

```
> mo = 10, mx = 10.2, sd = 2.25, n = 400
> zcal = (mx-mo) / (sd / sqrt(n))
> cat ("calculated value at z is = ", zcal)
calculated value at z is = 1.777778
> pvalue
[1] 0.07544036
```

value of p is greater than 0.05

we accept $H_0 : \mu = 10$

] 11.00 >
variance = (n-1) * var(x) / n
variance
10.59482
sd = sqrt (variance)
sd
3.254484
mo = 12.5
t = (mx-mo) / (sd / sqrt (n))
t
pvalue =12 * (1- pnorm (abs(t)))
pvalue
0.
>

2. In a random sample of 1000 students of us asked found that 750 use blue. Test the hypothesis that the population proportion is 0.8 at 1% level of significance.

=> H0 : P = 0.8 against H1 : P ≠ 0.8

> P = 0.8 ; q = 1-P ; p = 750/1000 ; n = ...

> zcal = (p-P) / sqrt (P * (q/n)))

> zcal cat ("calculated value of z: ... zcal)

calculated value of z: -3.952847

> pval = 2 * (1 - pnorm (abs (zcal)))

> cat ("calculated pvalue: ...

calculated pvalue: 7.72268e-05

∴ pvalue is less than 0.1, the value is rejected.

3. Two random samples of size 1000 and 2000 are drawn from two population with their means are 67.5 and 68. Test the hypothesis H0 : μ1 = μ2 as H1 μ1 ≠ μ2 at 5% level of significance.

H0 : μ1 = μ2 as H1 : μ1 ≠ μ2.

n1 = 1000 ; n2 = 2000 ; mx1 = 67.5 ; mx2 = 68 ; ...

zcal = (mx1 - mx2) / (sqrt ((sd ^ 2 / n1 ...

cat (" calculated value of z : ", zcal)

calculated value of z : -5.163978

pval = 2 * (1 - pnorm (abs (zcal)))

cat (" calculated pvalue : ", pval)

... pvalue is less than 0.05 then ...

Practical 6 :
Topic : large sample test

let the population mean (the amount spend per popu customer in a restraunt is 250 a sample of 100 customers selected the sample mean is calculated as 2~~7~~0 275 and S.D 30. test the hypothesis that population mean is 250 or not at 5% level of significance.

$H_0 : \mu = 275$ against $H_1 : \mu \neq 275$

```
> mx = 275
> mo = 250
> sd = 30
> n = 100
> zcal = (mx - mo) / (sd / sqrt (n))
> cat (" calculated value of z : ", zcal)
calculated value of z : 8.3333
> pvalue = 2 * (1 - pnorm (abs (zcal)))
> cat (" calculated pvalue : ", pvalue)
calculated pvalue : 0
```

pvalue is less than 0.5 then the value P is rejected.

The study of noise level in two hospitals is given below test the claim that the two hospitals have same noise at 1% level of significance

	Hospital A	Hospital B
size	84	34
mean	61.2	59.4
S.D	7.9	7.5

H0 : μ1 = μ2 * H2 : μ1 ≠ μ2

> n1 = 84
> n2 = 34
> mx1 = 61.2
> mx2 = 59.4
> sd1 = 7.9
> sd2 = 7.5
> zcal = (mx1 - mx2) / (sqrt (sd1^2 /n1) + (sd2^
> cat (" calculated value of z: ", zcal)
> calculated value of z: 1.1625528
> pval = 2 * (1 - pnorm (abs (zcal)))
> cat (" calculated Pvalue : ", pval)
calculated pvalue : ~~0.245023~~ 1.178520e-10
0.245023

Scanned with CamScanner

Practical -7 :

Topic : small sample test

The marks of 10 students are given by 63, 65, 66, 67, 68, 69, 70, 71, 72. Test the hypothesis that the sample comes from operation that the sample with average 66

=> H0 . μ = 66
> x = c (63, 63, 66, 67, 68, 69, 70, 71, 72)
> t . test (x)

one sample t - test

data : x
t = 68 . 319 , df = 9, p - value = 1.558 e-4
alternative hypothesis : true mean is not equal to 0 95 percent confidence interval
5 . 65171 to 0 . 14829
sample estimates :
mean of x
67 . 9
pvalue = 1.558e-13 < 0.05 . we reject i
> p value = 1.558 e -15
> it (pvalue > 0.05) { (cat (" accept H0")
else { cat (" Reject H0") }
Reject H0 .

The water data of 6 shops before and after a special campaign are given below. Test the hypothesis

Before : 55, 23, 31, 45, 50, 42
After : 58, 23, 30, 55, 56, 45

H0 - There is no significant difference before and after the campaign

> x = c (33, 28, 31 -48, 50, 42)
> y = c(58, 23, 50, 55, 56, 45,
> t test (x, y, paired = T, alternative = "greater")

paired t - test

data : x and y
t = -2.5815, df = 5, p-value = 0.9166
alternative hypothesis : true difference in means is greater than 0
35 percent confidence intervals

sample estimates:
mean of difference
-3.5

= 0.5606
> pvalue
> if (pvalue > 0.05) { cat ("Accept H0" } else
{ cat ("Reject H0 } Accept H0.

780

2) TU s groups of students score the following marks . Test the hypothesis that there is no significant difference between two groups.

group group 1 : 18 , 22, 21 , 17, 20, 17, 23, 20, 2
group 2 : 16, 20, 14 , 21, 20, 18 , 13, 1

=> H0 : There is no difference betw two group.

> x = c(18, 22, 21, 17, 20, 17, 23, 20, 22,
> y = c(16, 20, 14, 21, 20, 18 , 13, 15
> t. test (x, y)

Welch two sample t-

220

wo medians are applied to two
group of patients respectively
group 1 : 10, 12, 13, 11, 14
8, 9, 12, 14, 15, 10, 9

there any significancy difference
ween two medicines.
there is no significance difference
= c(10, 12, 13, 11, 14)
= c(8, 9, 12, 14, 15, 10, 9)
test (x, y)

which two sample t-test

x and y, dt = 9.794, p-value = 0.414
0.8030
native hypothesis
cent confidence interval
171 3.78171

n of x estimate mean of y
11

= 0.406.

p-value > 0.5) { cat("Accept H0") }
{reject H0} }

Practical 8 :

Topic : large and small sample tes

Ho : μ = 55 against H1 : μ ≠ 55

```
> mx = 55
> mo = 52
> n = 100
> sd = 7
> zcal = (mx-mo) / (sd / sqrt(n))
> zcal
[1] 4.285714
> pval = 2 * (1-pnorm (abs (zcal)))
> pval
[1] 1.82153e -05
```

(pval is rejected)

Ho : P = 0.5 against H1 : P ≠ 0.5

```
> P = 350/700
> p = 0.5
> n = 700
> q = 1 - p
> zcal = (P- p) / sqrt (p * (q/n)))
> zcal
[1] 0
> pval = 2 * (1 - pnorm (abs (zcal)))
> pval
[1] 1
```

(pval is accepted)

> pval = 2 * (1 - pnorm(abs(zcal)))
> pval
[1] 0.0124933 pval is rejected against H1 : μ ≠

Q5) H0 : μ = 66

> x = c(63, 63, 68, 69, 71, 71, ...
> t.test(x)

one sample t-test

data : x

t = 4.784, df = 6, p-value = 5.58...
alternative hypothesis the mean is not equal to 0. 95 percent confidence interval
64.66749 76.62992

sample alternative.
mean of x
68.42386

pval is rejected against H1 : $\sigma_1 \neq \sigma_2$

H0 : $\sigma_1 = \sigma_2$

> x = c(66, 67, 75, 76, 82, 84, 88 ...
> y = c(69, 66, 74, 78, 32, 35 ...
> var.test(x, y)
> var.test to compare two variances

Scanned with CamScanner

Practical - 09

Aim : chisquare tests and pin. VA

Q.1) use the following data, test centre the calculation of the name and child are independent or not

condition of name		clean	dirty
child condition	clean	70	50
	fairly clean	80	20
	dirty	35	45

Test the hypothesis that vaccine and disease are independent or not]

		vaccine	
		Affected	not affected
disease	Affected	70	46
	Not affected	35	37

850

perform a ANOVA for the following data a observations.

Type	50, 52
A	53, 55, 53
3	60, 58, 57, 56
	52, 54, 54, 55

250

: Practical 10 :

Aim : non. parametric test

Following are the amount of sulpher oxide the comi Hed by some industry in 20 days Apply sign test to test the hypothesis that the population median is 21.5 at 5% level of significance

data- 17, 15, 20, 20, 10, 10, 22, 25, 27, 9, 20, 17, 6, 24, 14, 15, 23, 24, 26

: Population median is 21.5.

= c (data)

e = 21.5

p = length (x [x > m e])

= length (x [x < me])

= sp + sn

= pbinom (sp, n, 0.5)

V =

V

0.411301

The following are values of sample
test the hypothesis that population
median is 60 against the alternative
is more than
ing
data: 63, 65, 60, 01, 89, 51, 68, 62, 63, 3,
69, 72, 69, 48, 66, 72, 62, 08, 37, 69.

: population median = 60, H1: population
median > 60

x = c(data)

wilcox.test (x, alter = "greater")

data: x

= 145, p-value = 0.0229

alternative hypothesis: true location
greater than 60

value < 0.05 (rejected)

of alternative is less: alter = "less"

If alternative is not equal
alter = "two sided"

If the alternative ~~is~~ greater;
alter = "greater"

g WSRT test the population median
2 or less than 12
7, 24, 25, 20, 21, 32, 28, 12, 25, 24, 26
median is 12, H1: population

> x = c(weight before)
> y = c(weight after)
> d = x-y
> wilcox.test (d, alter= "two.sided", mu = 0)

wilcoxon signed rank test with continuity correction

data : d
v = 4.5, p-value = 0.4981
alternative hypothesis : true location is not equal to 0

p value > 0.05 (accepted).

AM 5.3.20