THAKUR COLLEGE OF SCIENCE & COMMERCE

NAAC
Accredited
with Grade "A"
(3rd Cycle)

THAKUR
TRUSTS

ISO
9001 : 2015
Certified

Degree College

Computer Journal

CERTIFICATE

SEMESTER II UID No.

Class FY BSC - CS Roll No. 1864 Year 2019 - 2020

This is to certify that the work entered in this journal is the work of Mst. / Ms. Akshatha Swamy

9	Anova

RSO

PRACTICAL-1

Topic – Random Variable

Q1] Find the mean & variance for the following

X	-1	0	1	2
P(X)	0.1	0.2	0.3	0.4

Sol –

X	P(X)	X P(X)	$E(X)^2$	$[E(X)]^2$
-1	0.1	-0.1	0.1	0.01
0	0.2	0	0	0
1	0.3	0.3	0.3	0.09
2	0.4	0.8	0.16	0.64
TOTAL	$\Sigma_1 = 1$	$\Sigma_1 = 1$	$\Sigma_1 = 0.20$	$\Sigma_1 = 0.74$

$\therefore$ Mean $= E(x) = \Sigma x_i \cdot p(x) = 1$

$\therefore$ Variance $= V(x) = \Sigma_1 E(x)^2 - \Sigma_1 [E(x)]^2$

$= 2 - 0.74$

$= 1.24$

ANS Mean is 1 & Variance $\rightarrow$ 1.24

b]

X	-1	0	1	2
P(X)	1/8	1/8	1/4	1/2

Soln

X	P(X)	X·P(X)	$E(X)^2$	$[E(X)]^2$
-1	1/8	-1/8	1/8	1/64
0	1/8	0	0	0
1	1/4	1/4	1/4	1/16
2	1/2	1	2	1
TOTAL	$\Sigma = 1$	$\Sigma_1 = 9/8$	$\Sigma_1 = 19/8$	$\Sigma_1 = 69/64$

$\therefore$ Mean $= E(X) = \sum X \cdot P(X) = 9/8$

Variance $= V(X) = \sum E(X)^2 - \sum [E(X)]^2$

$= \frac{19}{8} - \frac{69}{64}$

$= \frac{83}{64}$

Ans: Mean is $9/8$ & Variance is $83/64$

c]

X	-3	10	15
P(X)	0.4	0.35	0.25

Soln

X	P(X)	X·P(X)	$E(X)^2$	$[E(X)]^2$
-3	0.4	-1.2	3.6	1.44
10	0.35	3.5	3.5	12.25
15	0.25	3.75	56.25	14.0625
TOTAL	$\sum = 1$	$\sum = 6.05$	$\sum = 94.85$	$\sum = 27.752$

$\therefore$ Mean $= E(X) = \sum X \cdot P(X) = 6.05$

Variance $= V(X) = \sum E(X)^2 - \sum E(X)$

$= 94.85 - 27.752$

$= 67.0975$

Ans:- Mean $\rightarrow$ 6.05 & variance $\rightarrow$ 67.0975

Q3) The p.m.f. of random variable X is given by

X	-3	-1	0	1	2	3	5	8
P(X)	0.1	0.2	0.18	0.2	0.1	0.15	0.05	0.05

Obtain cdf Find ① $P(-1 \le x \le 2)$ ② $P(16 \le x \le 25)$
③ $P(X \le 2)$ ④ $P(X \ge 0)$.

Soln -

X	-3	-1	0	1	2	3	5	8
P(X)	0.1	0.2	0.15	0.2	0.1	0.15	0.05	0.05
F(X)	0.1	0.3	0.45	0.65	0.75	0.90	0.95	1.0

① $P(-1 \le x \le 2) = P(X \le 2) - P(X \le -1) + P(X = -1)$
$= F(X_0) - F(X_a) + P(a)$
$= F(2) - F(-1) + P(-1)$
$= 0.75 - 0.3 + 0.2$
$= 0.25$

② $P(1 \le X \le 5) = F(X_b) - F(X_a) + P(a)$
$= F(5) - F(1) + P(1)$
$= 0.95 - 0.65 + 0.2$

Q4] Let f be continuous random variable with pdf

$$f(x) = \frac{x+1}{2} \quad -1 < x < 1$$
$$= 0 \quad \text{otherwise}$$

Obtain cdf of x. ~~Find mean &~~

Soln - By defn of cdf we have,

$$F(X) = \int_{0-1}^{+1} t\,dt$$

$$= \int_{0-1}^{1} \frac{x+1}{2}\,dx$$

$$= \int_{-r}^{1} \frac{1}{2}\left(\frac{1}{2}x^2 + x\right) \text{ for } -1 < x < 1$$

Hence the cdf is

$$F(X) = 0 \text{ for } x \leq -1$$
$$= \frac{1}{4}x^2 + \frac{1}{2}x \text{ for } -1 \leq x \leq 1$$
$$= 0 \text{ for } x \geq 1$$

Q5] Let f be continuous random variable with pdf

$$f(x) = \frac{x+2}{18} \quad -2 \leq x \leq 4$$
$$= 0 \quad \text{otherwise}$$

Calculate cdf.

Soln - By defn of cdf we have

$$F(X) = \int_{2}^{4} t\,dt$$

PRACTICAL-2

Topic - Binomial Distribution

Q1] An unbiased coin is tossed 4 times, calculate the probability of obtaining no head, at least one head & more than one tail.

a) No head -

```
> dbinom (0, 4, 0.5)
↳ 0.0625
```

b) At least one head

```
> 1- dbinom(0, 4, 0.5)
↳ 0.9375
```

c) More than one tail

```
> pbinom (1, 4, 0.5, lower.tail=F)
↳ 0.83692
```

Q2] The probability that student is accepted to a prestigious college is 0.3. If 5 students apply, what's the probability of atmost 2 are accepted.

```
> pbinom (2, 5, 0.3)
↳ 0.83692
```

Ho = $P_1 = P_2$
H_1 = $P_1 \neq P_2$

```
> (0.22 * 200 + 0.15 * 200 / (1/200 + 1 /200))
> 0.185
> 1-0.185
> 0.815
> z = sqrt (0.185 * 0.815 * (1/200 + 1/200))
> 2 * (1 - pnorm (abs (z)))
[1] 0.9969018
```

Accept Ho

Q. EQUALITY OF 2 POPULATION PROPORTION

1] In an early election campaign a telephone poll of 800 registered voter shows favor 460. In second poll opinion 520 of 1000 registered voter favored the candidate at 5% level of conf. Is there sufficient evidence that popularity has decreased

$H_0 = P_1 = P_2$
$H_1 = P_1 \neq P_2$

> P = (460 * 800 + 520 * 1000) / (520 + 1000)
> P
[1] .574
> 1 - 0.544
> z = sqrt (0.544 * 0.456 * (1/520 + 1/1000))
> z * (1 - pnorm(abs(z)))
[1] 0.5444
Accept H_0

Q2] From a ~~con~~ consignment 200 articles are drawn & 44 was found defective from consignment B 200 samples are drawn out of which 30 was found defective test whether the proportion of defective items in 2 consignment are significantly different

Q7]	Graduate	Undergraduate
online face	20	25
to face	40	8

Is there any association b/w student's preference for types of education & method

∴ H_0 = Independent, H_1 = Dependent.

> x = c(20, 40, 25, 8)
> z = matrix(x, nrow = 2)
> chisq.test(z)

Pearson's chi-sq test with Yates continuity correction

data: z

X-squared = 18.05, df = 1, p-value = 2.15e-5

∴ Reject null hypothesis

∴ Both are dependent.

$$= \int_2^4 \frac{x+2}{18}\, dx = \frac{1}{18}\left(\frac{1}{2} x^2 + 2x\right)$$

for $-2 \le x \le 4$

Hence cdf is

$F(X) = 0$ for $X < -2$

$= \frac{1}{18}\left(\frac{1}{2} x^2 + 2x\right)$

for $-2 < x < 4$

$= 0$ for $x \geq 4$

Klen

Scanned with CamScanner

PRACTICAL- 4
Testing of hypothesis

Q1] Sample mean & deviation given single population
A.1] Suppose the food level on the cook states that it has atmost 2g of saturated fats in a store cookies. In a sample of 35 cookies, it was found that mean amount of saturated fat per cookies 2.1g Assume that sample is 0.3 dt (5% level of conf) can be rejected the claim on food label

Ho = u < 2 , H1 = u > 2
>z = (2.1 - 2)/(0.3 /sqrt(35))
[1] 972027
>1 - pnorm(z)
[1] 0.0243
∴ Reject the null hypothesis.
∴ Accept H

Q2] A sample of 100 customers was randomly selected & it was found that avg spending was 275/- The SD 30. Using 0.05 level of significance would you conclude that amount spent by customer is more than 250/-

Ho = u < 250 , H1 = u > 250
> z = (275 - 250)/(30 /sqrt(100))
> [1] 8.333
> 1 - pnorm (z, 99) [1] 2.3057
∴ Reject Null hypothesis
∴ Accept H

Training

H_1 = no change in IQ
H_0 = IQ increased after training

```
> a = c(120, 118, 125, 136, 121)
> b = c( 110, 120, 123, 132, 125)
> z = sum ([b-a]^2/a)
> pchiq [z, df = length[b] - 1]
[1] 0.1135959
```

∴ Accept null hypothesis
∴ There is change in IQ after training.

3] In a big city, 325 men out of 600 men were found to be self employed conclusion is that maximum men in city are self-employed

$H_0 = \mu = 0.5$, $H_1 = \mu \neq 0.5$

> Z = (0.5 - 0.325) / sqrt((0.5 * 0.5)/600)

> Z * 1 - pnorm(abs(z))

[1] 0.04155239.

Reject H_0

Accept H_1

4] Experience shows that 20% of manufactures products are of top quality. In 1 day production of 400 articles only 50 are top quality. Test hypothesis that experience of 20% of manufactured is wrong.

$H_0 = \mu = 0.2$

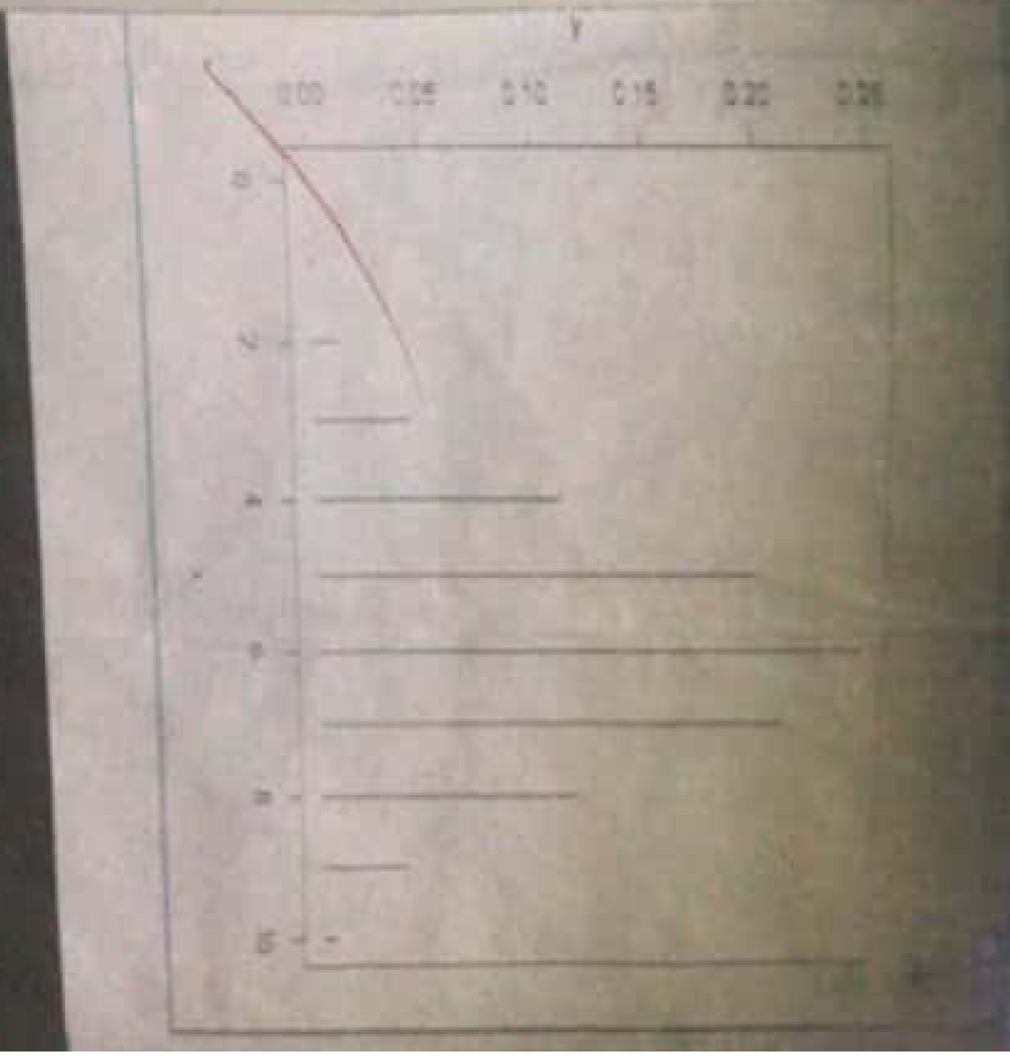
$H_1 = \mu_1 \neq 0.2$

> Z = (0.125 - 0.20) /sqrt((0.2 * 0.8)/400)

> Z * 1 - pnorm (abs(z))

[1] 0.0017.68346

Reject H_0, Accept H_1


y
0.00
0.05
0.10
0.15
0.20
0.25

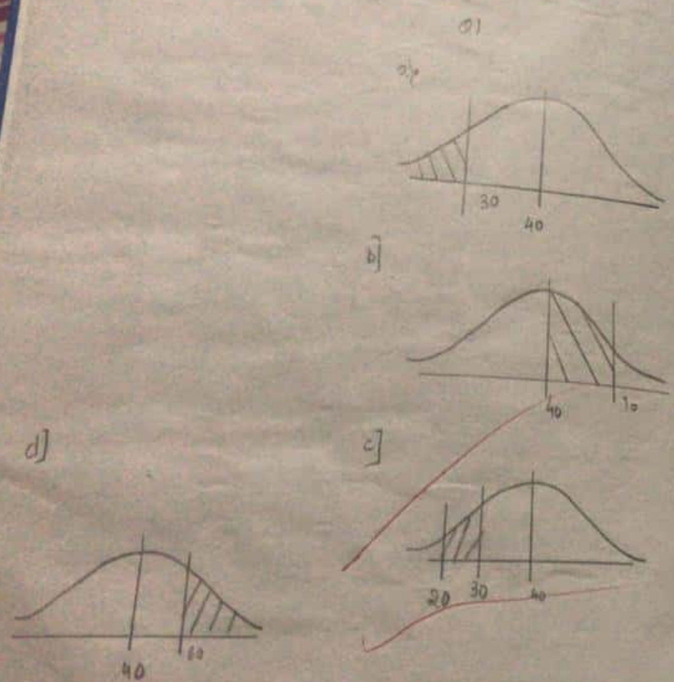
Q3 An unbiased coin is tossed 6 times the probability of head at any toss = 0·3. Let x be no of heads that comes up. Calculate P(X=2), P(X=3), P(1<x<5)

```
> dbinom (2, 6 ,0·3)
↳ 0·324135
> dbinom (3, 6, 0·3)
↳ 0·18522
> dbinom (2,6,0·3) + dbinom (3,6,0·3) + dbinom (4,6,0·3)
↳ 0·7437
```

Q4 For n = 10, p = 0·6, evaluate binomial probabilities and plot the graphs of pmf & cdf

```
> x = seq (0, 10)
> y = dbinom(x , 10, 0·6)
> y
[1]  0·0001048576   0·0015728640   0·0106168320
     0·0424673280   0·1114767360   0·2006581 2
     0·2508226560   0·2149908480   0·1209323950

> plot (x, y, xlab = "sequence", ylab = "probabilities", "o",
        pch = 16)
> x=seq(0,10)
> y = pbinom(x, 10, 0·6)
> plot (x, y, xlab = "sequence", ylab = "probabilities", "o",
        pch = 16).
```



PRACTICAL-03

Title - Normal Distribution

Q] A normal distribution of 100 students with mean = 40, SD = 15
Find no of students whose marks are
(1) P(X < 30) (2) P(40 < x < 70) (3) P(25 < x < 35)
(4) P(X > 60)

```
>pnorm(30, 40, 15)
↳ 0.25249 25

>pnorm(70, 40, 15) - pnorm(40, 40, 15)
↳ 0.4772499

> pnorm(38, 40, 15) - pnorm(25, 40, 15)
↳ 0.2107861

>1-pnorm(60, 40, 15)
↳ 0.09121122
```

Q] A quality control of engineer finds that sample of 100 light have avg life of 470 hours assuming population SD = 25 test whether the population mean is 480 hours. 1.05 → 0.05.

↳ Ho = μ < 480 , H₁ = μ > 480

> z = [470 - 480] / (25 /sqrt(100))

> [1] -4

> pt (z, 99, lower tail = T)

[1] 6.11257

Reject Null hypothesis

∴ Accept H₁.

Q] A principal at school claims that the IQ is 100 of the students. A random sample of 30 students whose IQ was found to 118. The SD of population = 15 test the claim of principal.

Ho = μ = 100

H₁ = μ > 100

> z = (112 - 100) / (15 / sqrt(30))

> [1] 4.38178

> pt (z, 99, lower tail = F)

[] 5.8856 , e = 06

∴ Reject Null Hypothesis

PRACTICAL-5

Chi-Square Test

Q1) Use the following data to test whether the alternative conditions of home & child are independent

		condition of home clean	Dirty
Condition of a child	clean	70	50
	fairly clean	80	20
	Dirty	35	45

H0 = Both are independent, H1 = Both are dependent

> x = c(70, 80, 35)

> y = c(50, 20, 45)

> z = data.frame(x, y)

[1]

#	x	y
1	70	50
2	80	20
3	35	45

> chisq.test(z)

Pearson's chi-squared test

data: z

x-squared = 25.646, df = 2, pvalue = 2.098, e

∴ Reject null hypothesis

∴ Both are dependent

Q2] If $P(X)$ is pmf of a random variable X. If $P(X)$ represents pmf for random variable X. Find value of K. Then evaluate mean & variance

Soln- As $P(X_i)$ is a pmf, it should satisfy the properties of pmf which are

i) $P(X_i) > 0$ for all sample space

ii) $\sum P(X_i) = 1$

X	-1	0	1	2
P(X)	K+1/13	K/13	1/13	K-4/13

$$\sum P(X) = 1 = \frac{K+1}{3} + \frac{K}{13} + \frac{1}{13} + \frac{K-4}{13}$$

$$1 = \frac{K+1 + K + 1 + K-4}{13}$$

$$13 = 3K - 2$$

$$15 = 3K$$

$$\boxed{K = 5}$$

X	P(X)	X·P(X)	E(X)²	[E(X)]²
-1	6/13	-6/13	6/13	36/189
0	5/13	0	0	0
1	1/13	1/13	1/13	1/189
2	1/13	2/13	4/13	4/189
TOTAL	$\sum = 1$	$\sum = -3/13$	$\sum = 11/13$	$\sum = 41/189$

$$\therefore \text{Mean} = \sum X \cdot P(X) = \frac{-3}{13}$$

$$\therefore \text{Variance} = \sum E(X)^2 - \sum [E(X)]^2$$

$$= \frac{11}{13} - \frac{41}{189} = \frac{102}{169}$$

Q4] A random variable X follows normal distribution with $\mu = 10$, $\sigma = 2$. Generate 100 obs & syntax & evaluate its mean, median & variance.

```
> x = rnorm(100, 10, 2)
> summary(x)
```

Min	1st Q	Median	Mean	3rd Q	Max
5.713	8.444	9.723	9.914	11.325	14.238

```
> var(x)
```
↳ 3.648924

Q5] Write a command to generate 10 r.n for normally distribution with $\mu = 50$, $\sigma = 4$. Find the sample mean & median.

```
> x = rnorm(10, 50, 4)
> summary(x)
```

Min	1st Q	Median	Mean	3rd Q	Max
44.73	50.46	52.01	52.35	54.39	58.85

Neat

SINGLE POPULATION PROPORTION

01] It is believed that coin is fair. The coin is tossed 40 times; 25 times head occurs. Indicate whether the coin is fair or not at 95%. po = 0.5, qo = 1 - po = 0.5

p = 28/40 = 0.7, n = 40, Ho = u = 0.5
H1 = u ≠ 0.5

```
> z = (0.7 - 0.5)/sqrt((0.5*0.5)/40))
> z
> 2 * (1-pnorm(abs(z))
[1] 0.0114 1204
```

Reject null hypothesis
Accept H1

2] In a hospital 400 females & 520 males & born in a week p. this confirm male & female are born equal in number.

Ho = u = 0.5, H = u1 ≠ 0.5

```
> z = (0.52 - 0.5)/(sqrt(0.5*0.5)/1000)
> [1] 1.26 45
> 2 * (1-pnorm(abs(z))
[1] 0.2060506
```

Reject Ho
Accept H1

049

Q 2] A dice is tossed 120 times & following results are obtained

No of terms	Frequency
1	30
2	25
3	18
4	10
5	22
6	15

Test the hypothesis that dice is unbiased

∴ Ho = dice is unbiased , H₁ = dice is biased

```
> obs = c(30, 25, 18, 10, 22, 15)
> exp = sum(obs)/length(obs)
> exp
[1] 20

> z = sum((obs - exp)^2/exp))
> pchisq (z, df = length [obs] - 1]
[1]  0.9566
```

∴ Accept null hypothesis

∴ Dice is unbiased.

Scanned with CamScanner

PRACTICAL-6

Topic - t-test

① Let x = 33.66, 33.37, 33.61, 34.14, 33.16, 33.57, 33.56, 33.76, 33.82, 33.55, 34.08, 34.01, 33.98, 34.29, 33.88, 33.74, 33.84, 33.74

Write the R command for following test hypothesis

① $H_0: \mu = 3400$, $H_1: \mu \neq 3400$

② $H_0: \mu = 3400$, $H_1: \mu > 3400$

③ $H_0: \mu = 3400$, $H_1: \mu < 3400$

at 95% level of confidence Also check at 95% level of confidence.

→ ① $H_0: \mu = 3400$

$H_1: \mu \neq 3400$

> x = c (33.66, 33.37, 33.61, 34.14, 33.16, 33.57, 33.56, 33.76, 33.82, 33.55, 34.08, 34.01, 33.98, 34.29, 33.88, 33.74, 33.84, 33.74)

> t.test (x, mu = 3400, alter = "two.sided", confidence.level = 0.95)

One sample t test

data : x

t = -4.4865, df = 10, p value = 0.0025 28.

95 percent level of confidence level.

3361.797 3386.03.

$\therefore$ Reject H_0
$\therefore$ Accept H_1

> t.test (x, mu= 3400, alter= "two sided", conf.level=
one sample test

data x:
t = -4.4865, df = 19, p-value = 0.0002528
~~stato~~ alternative hypo: two mean is not equal
to 3400
3360.33 3387.57
sample estimates:
mean of x:
3373.95

$\therefore$ Reject H_0 $\therefore$ Accept H_1

(e) $H_0 = \mu = 3400$
$H_1 = \mu > 3400$
> t.test = -4.4865 , df=19 p.value = 0.995
alternative hypothesis: ~~two~~ mean is greater than
3400

3363.91 Inf
sample ~~of~~ estimates:
Mean of x:
3373.95 $\therefore$ Accept H_0

Q5] A die is tossed 180 times

No of terms	frequency
1	20
2	30
3	35
4	40
5	12
6	43

Test the hypothesis that dice is unbiased

$H_0 \rightarrow$ dice is biased

$H_1 \rightarrow$ dice is unbiased.

> x = c(20, 30, 35, 40, 12, 43)

> chisq.test(x)

data: x

x-squared = 23.933, df = 5, p-value = 0.000

∴ Reject null hypothesis

∴ Dice unbiased

93 percent level of confidence.
Inf 3855 563
Sample estimates
Mean of x:
337 ~~83~~. 93
Reject H_0 Accept H_1

82] Below all the data of gain in weights on a different data A and B.
Diet A. ~~26~~, 25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25.
Diet B 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21
↳: $H_0 = a-b = 0$
$H_1 = a-b \neq 0$
>a= c(25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 25, 35)
>b = c(44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21)
t test (a, b, paired = T, alter = "two sided",
conf. level = 0.95)
Paired test
data : a and b
t = -0.62787, df = 11, p-value = 0.5429
alternative hypothesis true difference in means is not equal to 0
95 percent confidence interval
-14.2673 30 7.933947

Scanned with CamScanner

sample estimate
mean of difference
- 3.166667

∴ Accept H_0
∴ There is no difference in weights.

Q3] Eleven students gained the test after 1 month they again gave the test after the tutions, do the marks gives confidence that students have benefitted by coaching.

E_1 : 23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19
E_2 24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 18
test at 99 level of confidence!

$\hookrightarrow$ E_1 = c(23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19)
E_2 = c (24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 18)
∴ H_0: $e1 = e2$
H_1: $e1 < e2$

~~> t-test (e1, e2, paired= T, alter= "less", conf.level= 0.99)~~
~~paired t test~~
~~> data: e1 and e2~~
~~t = -1.4832, df = 10, p-value = 0.0844~~
~~alternative hypothesis: true difference in mean is less than 0~~

99 percent confidence Interval:
- Inf 0.863333

sample estimate
mean of difference
-1

∴ Accept H_0

t test (x, mu = 3400 ... 587
one sided t test alter = "greater", conf level = 0.97)
data : x
t = -4.4865, df = 19, p value = 0.9999 alter distance
3 347 ... less mean is greater than 3400
... 3 ... 57
sample estimate
mean of x :
3373.95
Accept Ho

(ii) Ho : $\mu_1 = 3400$
$H_1 : \mu_1 < 3400$
> t test (x, mu, 3400, alter = "less", conf level = 0.95
one sided t test)

data : x
t test -4.4865 df = 19, pvalue = 0.0001864
95 percent level of confidence Inf 338 304
~~Sample of x :~~
~~Mean of x :~~
3 373.95
Reject Ho Accept Ho .
> t test (x, mean = 3400, alter = "less", conf level = 0.9
data x
t = 4.4865 , df = 19, pvalue = 0.0001264
alternative hypothesis true mean is less than
3400

oa7

Q4) Two drugs for BP was given & data was collected b1 = 0.3, 1.6, 0.2, -1.2, 0.1, 3.4, 3.9, 0.8, 0.2.
b2 : 1.4, 0.8, 1.1, 0.1, 0.1, 4.4, 5.5, 1.6, 4.8, 3.4.

The two drugs have same of affect, check whether two drugs have same effect on patient or not

↳ $H_0 = d_1 = d_2$.
$H_1 = d_1 \neq d_2$

> d1 = c(0.7, -1.6, 0.2, -1.2, -0.1, 3.4, 3.7, 0.8
> d2 = c(1.9, 0.8, 1.1, 0.1, 0.1, 4.4, 5.5, 1.6, 4.6, 3.4).

> t.test (d1, d2, alter = "two.sided", paired = T, conf level = 0.95)

Paired t-test

data: d1 and d2

95 percent confidence interval

mean of the difference

-1.58

Reject H0

Accept H1

Q5 If there is difference in salaries for the same job in 2 different countries.

↳ CA = c(5300, 49950, 44366, 40470, 36963

CB = c(62990, 58850, 4995, 55263, 47674, 43532

Scanned with CamScanner

>H1 = S1 = S2
>H0 = S1 ≠ S2
> t.test (CA, Cb, paired = T, alter = "two sided"
conf.level = 0.95)

paired t test
data: Ca and Cb
t = -4.4569, df = 5, p value = 0.00666
alternative hypothesis: true ~~differ~~ of difference investments
not equal to 0
95 percent confidence interval:
-10404.821 -2792.848
Sample estimates:
Mean of the difference: -6598.863
∴ Reject H0
Accept H1

Nam

০২৫

PRACTICAL-7 F-Test

Q1 Life encepted in 10 regioned of India in 1990 & 2000 are given below. Test whether variance at the 2-times are the same
1990: 37, 39, 36, 42, 45, 44, 46, 49, 50, 51
2000: 44, 45, 47, 43, 42, 49, 50, 41, 48, 58, 42.

↳ x = c(37, 39, 42, 45, 46, 49, 50, 51)
y = c(44, 45, 47, 43, 42, 49, 50, 41, 48, 58, 42)
. var.test(x, y)
F test to compare two variances:
data: x and y.
F = 1.1449, numdf = 9, denom df = 11
pvalue = 0.81845
alternative hypothesis: true ratio of variances not equal to 1
95% confidence of interval:
0.3191005 4.478 0350
sample estimation.
Ratio of variances:
1.1449.

I	25 28 26 29 29 31 31 26 31
II	30 25 31 32 23 25 36 26 31 32 32 32 27 31

For following data test hypothesis for
(1) equality of two population mean (t test)
(2) equality of two proposition variance (F-test)

> x = c (175, 168, 145, 190, 81, 185, 175, 200)
> y = c (180, 170, 153, 130, 179, 183, 187, 005)
> var.test (x, y)

F test to compare two variance

data x and y.

F = 1.25, num df 7 denom df = 7,
p-value alternative hypothesis : the ratio of variances is not equal to 1
0.258289 0.2437398

sample estimates
ratio of variances.
1.250021

~~t.test (x)~~

~~one sample t-test.~~

~~data: 0~~

~~sample estimates~~
~~Mean of x :~~
~~177.375~~

Q3] The following are the prices of commodities in the sample of shop selected at random from different cities.

C·A cc (74.10, 77.70, 75.35, 74, 73.80, 79.30, 75.80, 76.80). 77.10, 76.40

C·B = c (70.80, 74.90, 76.20, 73.80, 78.10, 74.20, 69.80, 81.20.

4 var.test (CA, C·B).

F - test to compare two variances

data a and b.

F = 0.22579, num df = 9, denom df = 7,

pvalue = 0.04249

alternative hypothesis: true ratio of variances is not equal to 1

95 % confidence level:

sample estimate

Ratio of variance:

0.22579,

Import CSV file in excel, import file in R & apply the test to check equality of variance of 2 data

obs1 → 10, 15, 17, 11, 16, 20

obs2 → 15, 14, 16, 11, 12, 19

>Data = read.csv(file.choose(), header = TRUE)

> Data

	obs1	obs2
1	10	15
2	15	14
3	17	16
4	11	11
5	16	12
6	20	19

>attach(Data)

> var.test(obs1, obs2)

data: obs1 and obs2

F = 1.7068, num df = 5, denom df = 5,

p-value = 0.5717

alternative hypothesis: true ratio of varian is not equal to 1

95% ~~mean~~ confidence interval

0.238838 12.192860

sample estimates

1:30

Q5] I : 25, 28, 26, 28, 22, 29, 31, 31, 26, 31
II : 30, 25, 31, 32, 32, 23, 25, 31, 31, 32, 32, 32, 27, 31, 38, 24

at 95% of confidence level, check the ratio of population variation

$\therefore H_0 = \sigma_1^2 = \sigma_2^2$

$\therefore H_1 = \sigma_1^2 \neq \sigma_2^2$

```
> x = c(25, 28, 26, 28, 22, 29, 31, 31, 26, 31)
> y = c(30, 25, 31, 32, 32, 23, 25, 31, 31, 32,
    32, 32, 27, 31, 38, 24)
> var.test (x, y)
```

F test to compare two variance

p-value = 0.4535

$\therefore$ Accept H_0

$\therefore$ Variance of I & II are same.

Klub

PRACTICAL-8

The time of (in hrs) of 10 randomly selected 9 volt battery of a certain company is as follows 28.9, 15.2, 28.7, 72.5, 48.0, 52.4, 37.6, 49.5, 62.1, 54.1. Test the hypothesis that the population median is 63 against the alternative hypothesis less than 63 at 5% level of significance.

↳ a = c[28.9, 15.2, 28.7, 72.5, 48.0, 52.4,
37.6, 49.5, 62.1, 54.1)

↳ which (a > 63)

[4]

↳ y = which (a > 63)

↳ length (y)

↳ [1]

↳ d = which (a < 63)

↳ length (d)

[9]

↳ qbinom (0.05, 10, 0.5)

[12]

If qbinom value greater then d then accept the null hypothesis.

Q2 Following data gives weight of 40 students in random sample.

46, 49, 57, 64, 46, 67, 54, 48, 69, 61, 57, 54, 50, 48, 65, 61, 66, 54, 50, 48, 49, 62, 67, 49, 47, 55, 59, 63, 53, 56, 67, 49, 60, 64, 53, 50, 48, 51, 52, 54.

Use sign test to test whether mean weight of the population is 50 kg against alternative that its greater than 50 kg

↳ x = c (46, 49, 57, 64, 46, 67, 54, 48, 69, 61, 57, 54, 50, 48, 65, 61, 66, 54, 50, 48, 49, 62, 67, 49, 47, 55, 59, 63, 53, 56, 67, 49, 60, 64, 53, 50, 48, 51, 52, 54)

↳ Sp = which (x > 50)
↳ length (Sp)
[~~25~~ 12] [12]

↳ Sn = which (x < 50)
↳ length (Sn)
[12]

↳ qbinom (0.05, 40, 0.5)
[15]

So, Sp is greater than qbinom then Reject Null Hypothesis.

Median age of tourists in visiting a certain place is claimed to be 41 years. A random sample of 17 tourists have age 48, 52, 25, 29, 57, 39, 45, 36, 30, 49, 28, 39, 44, 63, 32, 65, 42.
Use sign test to check the claim

H_0 :- age = 41

H_1 :- age $\neq$ 41

```
y = c(48, 52, 25, 29, 57, 39, 45, 36, 30, 49, 28,
      39, 44, 63, 32, 65, 42)
z = which(y > 41)
length(z)
[1] 9
j = which(y < 41)
length(j)
[1] 8
qbinom(0.05, 17, 0.5)
[1] 5
```

Accept null hypothesis.

Q4. The time in minutes that a patient has to wait for consultation is recorded as follow
y = c(15, 17, 27, 25, 20, 21, 32, 28, 42, 25, 29, 26)
Use wilcox test to check whether the median in tai is greater than 5 % level of significance.

↳ H_0 :- mean ~~...~~ > 20

H_1 :- mean ~~...~~ < 20

↳ y1 = which(y > 50)

↳ z = which(y < 20)

↳ length(z)

[1] 3

wilcox.test (y, alternative = "greater")

[1] Wilcox signed rank test with continuity corrected.

data : x

V = 78 , p-value = 0.001253.

∴ Accept null hypothesis.

> aov (values~ind, e)
Terms :

	ind	Residuals
Sum of Squares	203.333	54.000
Deg. of Freedom	2	12

Residual Standard error : 2.12132

>oneway.test (values~ind, e)
data : values and ind
F = 21.537, num df = 2.000, denom df = 7.9314,
p-value = 0.0006232
Reject the null hypothesis.

The following gives the life of tayes of brand. Test the hypothesis.

A	B	C	D
20	18	21	15
23	15	19	14
18	17	22	16
17	20	17	18
22	16	20	14
24	17		16

```
> a = c (A)
> b = c (B)
> c = c (C)
> d = c (D)
> L = list (a1 = a, b1 = b, c1 = c, d1 = d)
> L
$ a1
[1]  20  23  18  17  22  24

$ b1
[1] 18  15  17  20  16  17
```

> o = c (data of no-exercise)
> p = c (data of 20min)
> q = c (data of 60min)
> L = list(a1 = o, b1 = p, c1 = q)
> L
$ [a1]
23 26 81 48 58 37 29 44
$ [b1]
22 27 29 39 46 48 49 65
$ [c1]
59 66 38 49 56 60 56 62
>e = stack(L)
>e
Data is arranged
>oneway test (values~ind, e)
data values and ind
F = 5.9169, num df = 2.000, denom df = 13.334
p-value = 0.01633
Reject Null hypothesis.

> a = c (data of A)
> b = c (data of B)
> c = c (data of c)
> L = list (a1 = a, b1 = b, c1 = c)
> L

$a1
[1] 44 45 46 47 48 49

$b1
[1] 40 42 50 52 55

$c1
[1] 50 53 58 59
> e = stack(L)
> e
The data is arranged in ~~contnow~~ continous form
> oneway.test (values ~ ind, e)
data: values and ind
F = 6.325, num df = 2.000, denom df = 5.413
p-value = 0.0382
∴ Reject the null hypothesis.

An experiment was conducted on 8 person & the observation noted are.

NO exercise - 23, 26, 51, 48, 58, 37, 29, 44
20 min - 22, 27, 29, 39, 46, 48, 49, 65
60 min - 59, 66, 38, 49, 56, 60, 56, 62
Test the hypothesis are equal.

150

$ c1
[1] 21 19 22 17 20

$ d1
[1] 15 14 16 18 14 16
> e = stack(L)
> e
The data is formed in the way of continous data.
> oneway.test(values~ind, e)

data: values and ind
F = 6.6498, num df = 3.000, denom df = 10.093
pvalue = 0.009376
∴ Reject the null hypothesis.

3] Three types of wax is applied for the protection of cars & no of days of protection were needed. Test them are equally effective

A	B	C
44	40	50
45	42	53
46	51	58
47	52	59
48	55	
49		

Test whether 3 are equally effective.

Scanned with CamScanner

PRACTICAL-9

Q1 The following data gives the effect of 3 treatments

T1	T2	T3
2	10	10
3	8	13
7	7	14
2	5	13
6	10	15

Test the hypothesis that all treatment are equally effected

```
>t1 = c(2, 3, 7, 2, 6)
>t2 = c(10, 8, 7, 5, 10)
>t3 = c(10, 13, 14, 13, 15)
>data = data.frame(t1, t2, t3)
>data
   t1  t2  t3
1   2  10  10
2   3   8  13
3   7   7  14
4   2   5  13
5   6  10  15.
>e = stack(data)
>e
```

The numbers are arranged in the form of data

The weight in kgs of the person before and after stop smoking is

Before: 65, 75, 75, 72, 62

After: 72, 72, 72, 66, 82.

use wilcox test to check weight of a person is less after smoking at 5% level of significance

↳ x = c (65, 75, 75, 72, 62)

y = c (72, 72, 72, 66, 82)

z = x - y

z

[1] -7 0 3 -7 -4

wilcox.test (z, mu = 0, alternative = "less")

correction on

data : z

V = 1, p-value = 0.09873

alternative hypothesis true location is less than 0

Reject null hypothesis.

Q2] If the random variable x follows the normal distribution with mean = 50, V = 100
And ① P(X < 70) ② P(X > 65) ③ P(X ≤ 32)
④ P(35 < x < 60) ⑤ P(20 < x < 30)

a) >pnorm(70, 50, 10)
↳ 0.9772499

b) >1-pnorm(65, 50, 10)
↳ 0.03593032

c) >pnorm(32, 50, 10)
↳ 0.0668072

d) > pnorm(60, 50, 10) - pnorm(35, 30, 10)
↳ 0.7745375

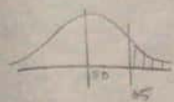
e) >pnorm(30, 50, 10) - pnorm(20, 50, 10)
↳ 0.02140023

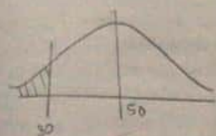
Q3] Let X~N(30, 50, 160, 400) find K1 & K2 such that P(X < K1) = 0.8 & P(X > K2) = 0.8

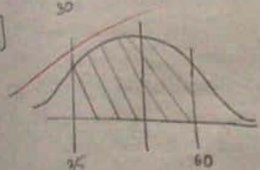
> qnorm(0.6, 160, 20)
↳ 165.0669
> qnorm(0.8, 160, 20)
↳ 176.8324

Q2]

a]


b]


c]


d]


e]