Practical 01

* Basics of R-software

R is a software for data analysis of statistical computing. Thus software is used for effective data handling & output storage is possible.
It is capable of graphical display.
It is a free software

1) $2^2 + \sqrt{25} + 35$

→ 2^2 + sqrt(25) + 35

= 44

2) $2 \times 5 \times 3 + 62 \div 5 + \sqrt{49}$

→ 2*5*3 + 62/5 + sqrt(49)

= 49.4.

3) $\sqrt{76 + 4 \times 2 + 9 \div 5}$

→ sqrt(76 + 4*2 + 9/5)

= 9.262829.

4) $42 + |-10| + 7^2 + 3 \times 9$

→ 42 + abs(-10) + 7^2 + 3*9

= 128.

Q. $x = 20$; $y = 30$

find, $x+y$; x^2+y^2 ; $\sqrt{y^3-x^3}$; abs(x−y)

i) $x+y$

= 50.

ii) x^2+y^2

→ x^2 + y^2

= 1300.

iii) $\sqrt{y^3-x^3}$

→ y^3 − x^3

= 137.8405

iv) abs(x − y)

= 10.

Q. calculate the following.

1) c(2,3,4,5)^2

→ 4 9 16 25.

2) c(4,5,6,8) * 3

→ 12 15 18 24.

3) c(2,3,5,7) * c(-2,-3,-5,-4)

→ -4 -9 -25 -28.

4) c(2,3,5,7) * c(8,9)

→ 16 27 40 63

5) c(1,2,3,4,5,6)^c(2,3)

→ 1 8 9 64 25 216.

2) Find the sum, prod, min, max of the given value
5, 8, 6, 7, 9, 10, 15, 5.

→ x = c(5, 8, 6, 7, 9, 10, 15, 5)

length(x)
= 8

sum(x)
= 65

max(x)
= 15

min(x)
= 5.

prod(x)
= 11340000.

Matrix calculation.

1) $\begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$.

→ X ← matrix (nrow = 4, ncol = 2, data = c(1,2,3,4,5,6,7,8))

2) $X' = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ $Y = \begin{bmatrix} 2 & 4 & 10 \\ -2 & 8 & -11 \\ 10 & 6 & 12 \end{bmatrix}$

Find $x + y$; $x * y$, $2x + 3y$

X ← matrix (nrow = 3, ncol = 3, data = c(1, 2, 3, 4, 5, 6, 7, 8, 9))

Y ← matrix (nrow = 3, ncol = 3, data = c(2, -2, 10, 4, 8, 6, 10, -11, 12))

6.2

i) $X + Y$

$$\begin{bmatrix} 3 & 8 & 17 \\ 0 & 15 & 3 \\ 18 & 12 & 21 \end{bmatrix}$$

ii) $X * Y$

$$\begin{bmatrix} 0 & 92 & 70 \\ 4 & 40 & 33 \\ 80 & 86 & 108 \end{bmatrix}$$

iii) $2 * X + 3 * Y$

$$\begin{bmatrix} 9 & 30 & 44 \\ -3 & 84 & 17 \\ 86 & 50 & 64 \end{bmatrix}$$

6 Perform the following.

```
y = c(2,4,6,1,3,5,7,18,15,14,17,10,3,3,2,5,10,15,9,
      14,18,10,12)
→ length(y)
  23
a = table(y)
transform(a)
```

X	Freq
0	1
1	1
2	2
3	3
4	1
5	2
6	1
7	1
9	1
10	1
12	1
14	2
15	1
16	1
17	1
18	2
19	1

```
breaks = seq(0,20,5)
b = cut(x, breaks, light = FALSE)
c = tabl(b)
transform(c)
```

b	Freq
[0,5)	8
[5,10)	5
[10,15)	4
[15,20)	6

Practical - 02.

Title : problems on P.D.F and C.D.F

can the following be P.D.F?

i) $F(x) = \begin{cases} 2-x & ; 1 \le x \le 2 \\ 0 & \text{O.W} \end{cases}$

To prove :- $\int F(x)\,dx = 1$.

$= \int_1^2 (2-x)\,dx$

$= \int_1^2 2\,dx - \int_1^2 x\,dx$

$= \left[2x\right]_1^2 \left[-\frac{x^2}{2}\right]_1^2$

$= (4-2) - (2-0.5)$

$\neq 1$.

ii) $F(x) = \begin{cases} 3x^2 & ; 0 < x < 1 \\ 0 & \text{O.W.} \end{cases}$

To prove :- $\int F(x)\,dx = 1$.

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

$= \int_0^1 \frac{3x^3}{3}$

$a^n = \frac{a^{n+1}}{n+1}$

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

$= \left[\frac{3}{3}\right] - 0$

$= 1$

Hence, it is P.D.F

iii) $f(x) = \begin{cases} \frac{3x}{2}\left(1-\frac{x}{2}\right) & ; 0 \le x \le 2 \\ 0 & \text{O.W} \end{cases}$

To prove :- $\int f(x)\,dx = 1$.

$= \int^2 \frac{3x}{2}\left(1-\frac{x}{2}\right) dx$

$= \int_0^2 \frac{3x}{2} - \frac{3x^2}{4}\,dx$

$= \left[\frac{3x^2}{4}\right]_0^2 - \left[\frac{3x^3}{12}\right]_0^2$

$= 3/11\ HIN3$

$= 3 - 2$

$= 1$

Hence, it is P.D.F

Can the following be P.M.F ?

(i)

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

$P(3) = -0.1$

Since, one probability is negative

Hence, it is not P.M.F

(ii)

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

Since $P(x) \geq 0 \ \forall x$

and $\Sigma P(x) = 1$.

Hence, it is a P.M.F.

(iii)

x	10	20	30	40	50
$P(x)$	0.2	0.3	0.3	0.2	0.2

$$\Sigma P(x) = 0.2 + 0.3 + 0.3 + 0.2 + 0.2$$
$$= 1.2$$

Since $\Sigma P(x) \neq 1$

$P(x)$ is not a P.m.f.

Q.4 Find c.d.f

(i)

x	0	1	2	3	4	5	6
$P(x)$	0.1	0.1	0.2	0.2	0.1	0.2	0.1

→ $F(x) = 0$, $x < 0$

$= 0.1$, $0 \le x < 1$

$= 0.2$, $1 \le x < 2$

$= 0.4$, $2 \le x < 3$

$= 0.6$, $3 \le x < 4$

$= 0.7$, $4 \le x < 5$

$= 0.9$, $5 \le x < 6$

$= 1.0$, $x \ge 6$.

(ii)

x	1.0	1.2	1.4	1.6	1.8
$P(x)$	0.2	0.35	0.15	0.25	0.1

$F(x) = 0.2$, $x < 10$

$= 0.55$, $10 \le x < 12$

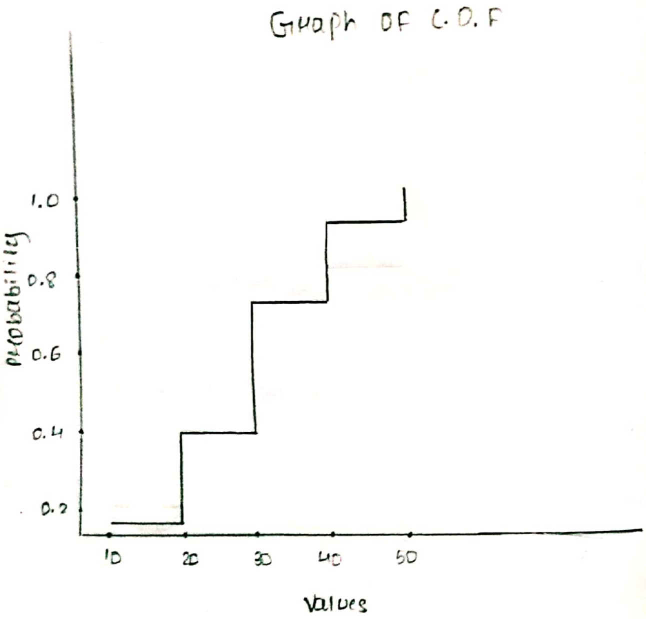
$= 0.70$, $12 \le x < 14$

$= 0.90$, $14 \le x < 16$

$= 1.0$, $x \ge 18$

M

Practical.



Practical No. 3

Title: Binomial Distributions and probability.

Q. Find the C.D.F. of the following P.D.F and draw the graph.

x	10	20	30	40	50
$P(x)$	0.15	0.25	0.3	0.2	0.1

$F(x) = 0$	if $x < 10$
$= 0.15$	$10 \leq x < 20$
$= 0.40$	$20 \leq x < 30$
$= 0.70$	$30 \leq x < 40$
$= 0.90$	$40 \leq x < 50$
$= 1.00$	$x \geq 50$

```
→ X = c(10,20 ,30 ,40,50)
  x
[1]  10  20  30  40  50
  PX = c(0.15, 0.25, 0.3 ,0.2,0.1)
  cumsum(PX)
[1]  0.15 0.40 0.70 0.90 1.00
  plot (x, cumsum(PX), xlab="Values", ylab="probability",
      Main = "Graph of C.D.F.", "s")
```

Q. Suppose there are 12 MCQ's in a test each question has 5 options & only one is correct. Find the probability of having:

1) five correct answers.
2) Atmost four correct answers.

→ It is given that
$n = 12$, $P = \frac{1}{5}$, $Q = \frac{4}{5}$.

x = Total no. of correct answers
$x \sim B(n, p)$

1) $n = 12$; $P = 1/5$; $q = 4/5$; $x = 5$
dbinom (5, 12, 1/5)
[1] 0.05315

2) $n = 12$; $P = 1/5$; $q = 4/5$; $x = 4$
pbinom (4, 12, 1/5)
[1] 0.9274

Q. There are 10 members in a committee. The probability of any member attending a meeting is 0.9, find the probability:

1) 7 members attended.
2) Atleast 5 members attended.
3) Atmost 6 members attended.

→ It is given that
$n = 10$, $P = 0.9$, $Q = 0.1$.

1) $n = 10$; $P = 0.9$; $Q = 0.1$; $X = 7$.
dbinom (7, 10, 0.9).
[1] 0.05739

2) $n = 10$; $P = 0.9$; $q = 0.1$; $X = 5$
1 - pbinom (5, 10, 0.9).
[1] 0.99836

3) $n = 10$; $P = 0.9$; $q = 0.1$; $X = 6$
pbinom (6, 10, 0.9)
[1] 0.01279

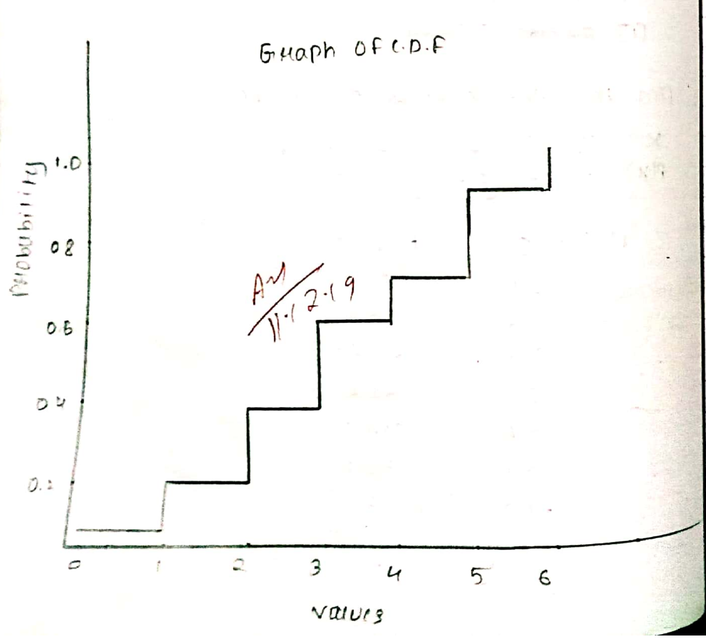
Q. Find the C.D.F & draw the graph.

x	0	1	2	3	4	5	6
$P(x)$	0.1	0.1	0.2	0.2	0.1	0.2	0.1

$F(x) = 0$ if $x < 0$
$= 0.1$ $0 \le x < 1$
$= 0.2$ $1 \le x < 2$
$= 0.4$ $2 \le x < 3$
$= 0.6$ $3 \le x < 4$
$= 0.7$ $4 \le x < 5$
$= 0.9$ $5 \le x < 6$
$= 1.0$ $x \ge 6$

110

```
→ x = c(0,1,2,3,4,5,6)
x
[1] 0 1 2 3 4 5 6
Px = c(0.1,0.1,0.2,0.2,0.1,0.2,0.1).
cumsum(Px).
[1] 0.1 0.2 0.4 0.6 0.7 0.9 1.0
Plot(x,cumsum(px), xlab="values", ylab=" Probability",
   main="Graph of C.D.F.", "s").
```



AM 11-12-19

Practical No. 4.

1. Find the complete binomial distribution
$n = 5$ & $P = 0.1$

2. Find the probability of exactly 10 success in 100 trial with $p = 0.1$.

3. X follows binomial distribution with $n = 10$, $P = 0.25$.
Find i) $P(X = 5)$.
ii) $P(X \leq 5)$
iii) $P(X > 7)$
iv) $P(5 < X < 7)$.

4. Probability of salesman makes a sales to the customer is 0.15. Find the probability
i) No sale for 10 customers
ii) More than 3 sale in 20 customer.

5. A student writes 5 Mcq. Each question has 4 options out of which 1 is correct. Calculate the probability for atleast 3 correct answers

6. X follows binomial distribution $n = 10$, $P = 0.4$ plot the graph PMF & CDF

1. $n = 5, p = 0.1$

→ dbinom (0:5, 5, 0.1)

[1] 0.59049 0.32805 0.07290 0.00810 0.00045 0.00001

2. $n = 100, x = 10, p = 0.1$

→ dbinom (10, 100, 0.1)

[1] 0.1318653

3. $n = 12, p = 0.25$

→ (i) $P(X = 5)$

dbinom (5, 12, 0.25)

[1] 0.1032414

(ii) $P(X \leq 5)$

pbinom (5, 12, 0.25)

[1] 0.9455978

(iii) $P(X > 7)$

$P(X > 7) = 1 - P(X \leq 7)$ = 1 - pbinom (7, 12, 0.5)

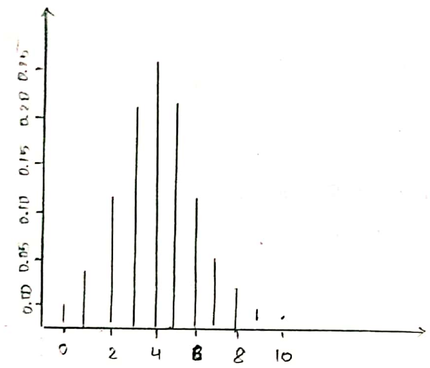
[1] 0.00278151

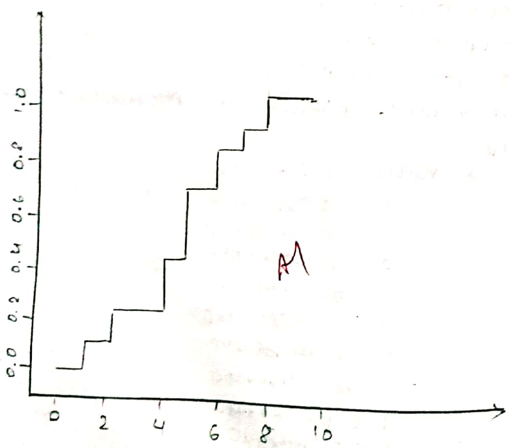
(iv) $P(5 < x < 7)$

dbinom (6, 12, 0.25)

[1] 0.00278151

plot(x, prob, "h")
plot(x, comprob, "s").





Practical No. 5.

Normal distribution.

1. $P[X = x] = \text{dnorm}(x, \mu, \sigma)$

2. $P(X \leq x) = \text{pnorm}(x, \mu, \sigma)$

3. $P(X > x) = 1 - \text{pnorm}(x, \mu, \sigma)$.

4. $P[x_1 < X < x_2] = \text{pnorm}(x_2, \mu, \sigma)$

5. To find the value of k so that :
$P(X \leq k) = P_1$; $\text{qnorm}(P_1, \mu, \sigma)$.

6. To generate 'n' random numbers :
$\text{rnorm}(n, \mu, \sigma)$.

Q.1 $X \sim N(\mu = 50, \sigma^2 = 100)$
Find i) $P(X \leq 40)$
ii) $P(X > 55)$
iii) $P(42 \leq X \leq 60)$
iv) $P(X \leq k) = 0.7$; $k = ?$

Q.2 $X \sim N(\mu = 100, \sigma^2 = \cancel{25})$.
i) $P(X \leq 110)$
ii) $P(X \leq 95)$
iii) $P(X > 115)$.
iv) $P(95 \leq x \leq 105)$
v) $P(X \leq k) = 0.4$; $k = ?$

Q.3. Generate 10 ~~more~~ random numbers from normal distribution with mean (μ) = 60, S.D = 5 also calculate the sample mean, median, variance & S.D.

Q.4. Draw the graph of standard normal distribution.

Q.1

→ (i) a = pnorm (40, 50, 10)

cat ("P(x<=40) is = ", a).

P(x<=40) is = 0.1586553 >

(ii) b = 1 - pnorm (55, 50, 10).

cat ("P(x>55) is = ", b).

P(x>55) is = 0.8085375 >.

(iii) c = pnorm (60, 50, 10) - pnorm (42, 50, 10)

cat ("P(42<=x<=60) is = ", c).

P(42<=x<=60) is = 0.6294893 >.

(iv) d = qnorm (0.7, 50, 10)

cat ("P(x<=k) = 0.7, k is = ", d).

P(x<=k) = 0.7, k is = 55.244017 >.

Q.2.

→ (i) a = pnorm (110, 100, 6)

cat ("P(x<=110) is = ", a).

P(x<=110) is = 0.9522096 >.

(ii) b = pnorm (95, 100, 6).

cat ("P(x<=95) is = ", b).

P(x<=95) is = 0.2023284 >.

(iii) c = 1 - pnorm (115, 100, 6).

cat ("P(x>115) is = ", c).

P(x>115) is = 0.006209665 >

(iv) d = pnorm (95, 100, 6) - pnorm (105, 100, 6).

cat ("P(95<=x<=105) is = ", d).

P(95<=x<=105) is = -0.5953482 >.

(v) e = qnorm (0.4, 100, 6).

cat ("P(x<=k) = 0.4, k is = ", e).

P(x<=k) = 0.4, k is = 98.47992 >.

PP.

Q.3.

→ n = 10, μ = 10, σ = 5

x = rnorm (10,60,5).

[1] 66.05727 51.97402 59.97334 60.87027 58.27801
56.90084 63.89487 56.54952

[9] 58.59008 65.11264.

am = mean (x).

>am

[1] 59.81959.

me = median(x).

> me

[1] 59.81959.

variance = (10-1) * var(x)/10.

> variance

[1] 16.93294.

sd = sqrt (variance).

>sd

[1] 4.114966.

Practical No. 6.

TOPIC :- Z distribution.

* Mean Test :

Q.1 Test the Hypothesis (H_0):

(i) $H_0 : \mu = 10$ against $H_1 : \mu \neq 10$.

A sample of size 400 is selected which gives a mean 10.2 & standard deviation 2.25. Test the hypothesis at 5% level of significance.

→

mo (mean of population) = 10

mx (mean of sample) = 10.2

sd = 2.25

n (sample size) = 400.

zcal = (mx − mo) / (sd / (sqrt(n)))

cat = ("zcal is :", zcal)

>> zcal is 1.77778.

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

pvalue

>> 0.07544 , Thus hypothesis accepted.

(NOTE : If result of tested hypothesis is > 0.05, then assumed hypothesis of $H_0 : \mu = 10$ is accepted as verified / correct).

Q.2 Test the hypothesis (H_0):

$H_0 : \mu = 75$ against $H_0 : \mu \neq 75$

A sample of size 100 is selected & the sample mean is 80, with SD of 3. Test the hypothesis at 5% level of significance.

→ mo = 75

mx = 80

sd = 3

n = 100

zcal = (mx − mo) / (sd / (sqrt(n)))

cat = ("zcal : ", zcal)

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

cat (pvalue is :", pvalue)

OUTPUT :

>> zcal : 16.667

>> pvalue is : 0

Thus, hypothesis rejected.

Q3) Test the Hypothesis (H_0) :-

$H_0 : \mu = 25$ against H_0 $\mu \neq 25$

A sample of 30 is selected. Test the hypothesis at 5% level of significance.

sample (x) = 20, 24, 27, 35, 30, 46, 26, 27, 10, 20, 30, 37, 35, 21, 22, 23, 24, 25, 26, 27, 28, 24, 30, 39, 27, 15, 19, 22, 20, 18.

→ mx = mean(x)

mx = 26.66

n = length (n)

variance = (n-1) * var (x) / n

sd1 = sqrt (variance)

>> sd = 2.2798

mo = 25

zcal = (mx - mo) / (sd / sqrt(n))
>> zcal = 0.8025
pvalue = 2 * (1 - pnorm (abs (zcal)))
>> pvalue = 0.4223.
Thus, hypothesis accepted.

* Proportion Test:

Q.1) Experience has shown that 20% students of a college smoke. A sample of 400 students reveal that out of 400 only 50 smoke. Test the hypothesis that the experience kept the correct proportion or not.
→ H0 pp = 0.2 (20%)
q = 1 - pp
sp = 50 / 400
n = 400

zcal = (sp - pp) / (sqrt (pp * q / n))
>>> zcal = -3.75
Pvalue = 2 * (1 - pnorm (abs (zcal)))
>>> Pvalue = 0.00017.
Thus, hypothesis is rejected.

Q.2 Test the hypothesis: H0: μ = 0.5 against p ≠ 0.5. A sample of 200 in selected where sp = 0.56 & pp = 0.5, level of significance is 1%.
→ H0 pp = 0.5
q = 1 - pp
sp = 0.56
n = 200
zcal = (sp - pp) / (sqrt (pp * q / n))
>>> zcal = 1.6970
pvalue = 2 * (1 - pnorm (abs (zcal)))
>> pvalue = 0.08968.

Thus, hypothesis is accepted.

54

Practical No. 7

Topic :- Large sample Tests.

Q.1 H_0: A study of noise level in two hospitals is calculated below, Test the hypothesis that the noise level in two hospitals are same or not.

	HOS. A	HOS. B
No. of sample obs	84	34
Mean	61	59
S.D	7	8

→ H_0: The noise levels are same.

n1 = 84

n2 = 34.

mx = 61

my = 59

sdx = 7

sdy = 8.

z = (mx - my)/sqrt((sdx^2/n1) + (sdy^2/n2))

[1] 1.273682.

cat ("z calculated is =", z)

{ z calculated is = ~~1.~~ 1.273682 >

pvalue = 2 * (1 - pnorm (abs(z)))

[1] 0.2027761.

Since pvalue > 0.05, we accept ~~reject~~ H_0 at 5% level of significance.

Q.2 H_0: Two random samples of size 1000 & 2000 are drawn from population with the means 67.5 & 68 respectively & with the same standard deviation of 2.5. Test the hypothesis that the mean of two population are equal.

→ H_0: Two populations are equal.

n1 = 1000

n2 = 2000

mx = 67.5

my = 68

sdx = 2.5

sdy = 2.5

z = (mx - my)/sqrt((sdx^2/n1) + (sdy^2/n2))

z

[1] -5.163978

cat ("z calculated is =", z)

z calculated is = -5.163978 >

pvalue = 2 * (1 - pnorm (abs(z)))

pvalue

[1] 2.417564e-07

Since, pvalue < 0.05, we reject the H_0 at 5% level of significance.

81

Q.3 H_0: In FYBSC, 20% of a random sample of 400 students had defective eye sight. In SYBSC, 15.5% of 500 students had the same defect. Is the difference of proportion same?

→ H_0: The proportion of the population are equal.

n1 = 400
n2 = 500
p1 = 0.2
p2 = 0.155
p = (n1 * p1 + n2 * p2) / (n1 + n2)
[1] 0.175
q = 1 - p
[1] 0.825
z = (p1 - p2)/sqrt (p * q * (1/n1 + 1/n2))
[1] 1.76547
cat ("z calculated is = ", z)
z calculated is = 1.76547 >
pvalue = 2 * (1 - pnorm (abs (z)))
pvalue
[1] 0.07748427

Since, pvalue > 0.05, we accept the H_0, at 5% level of significance.

Q.4 H_0: From each of the box of the apples, a sample size of 200 is collected. It is found that there are 44 bad apples in the 1st sample & 30 bad apples in the 2nd sample, Test the hypothesis, that the two boxes are equivalent in terms of no. of bad apples.

→ H_0: The two boxes are equivalent.

n1 = 200
n2 = 200
p1 = 44/200
p2 = 30/200
p = (n1 * p1 + n2 * p2)/(n1 + n2)
[1] 0.185
q = 1 - p
[1] 0.815
z = (p1 - p2) / sqrt (p * q * (1/n1 + 1/n2))
[1] 1.802741
pvalue = 2 * (1 - pnorm (abs (z)))
pvalue
[1]. 0.07142888

Since, pvalue > 0.05, we accept the H_0, at 5% level of significance.

Q.5 H_0: In MA class, out of a sample of 60, mean height is 63.5 inch with the SD 2.5. In a M.COM class out of 50 students, mean height is 69.5 inch with the SD as 2.5. Test the hypothesis that the mean of MA and Mcom class are same.

→ H_0: Heights of two classes is same.

n1 = 60

n2 = 50

mx = 63.5

my = 69.5

sdx = 2.5

sdy = 2.5

z = (mx-my)/sqrt((sdx^2)/n1)+(sdy^2/n2))

[1] 145.1808

pvalue = 2*(1-pnorm(abs(z)))

[1] 0

Since, pvalue < 0.05 we reject the H_0, at 5% level of significance

A-1

Practical No. 8.

Topic :- Small sample Test.

Q1 The flower's stems are selected & the heights are found to be 63, 63, 68, 69, 71, 71, 72 cm. Test the hypothesis that the mean height is 66 cm or not at 1% level of significance.

→ H_0: Mean = 66 cms.

x = c(63, 63, 68, 69, 71, 71, 72)

x

[1] 63 63 68 69 71 71 72

t.test(x)

One sample t-test

data: x

t = 47.94, df = 6, p-value = 5.522e-09

alternative hypothesis: true mean is not equal to 0

95% confidence interval:

64.66479 71.62092

sample estimates:

mean of x

68.14286

Since, p-value < 0.01, we reject the H_0 at 1% level of significance.

118

Q.2 Two random samples were drawn from two different population.

Sample 1 : 8, 10, 12, 11, 16, 15, 18, 7.

Sample 2 : 20, 15, 18, 9, 8, 10, 11, 12.

Test the hypothesis that there is no difference between the two population mean at 5% level of significance.

→ H_0 : There is no difference in the population means.

```
X = c(8, 10, 12, 11, 16, 15, 18, 7)
Y = c(20, 15, 18, 9, 8, 10, 11, 12)
t.test(X, Y)
```

Welch two sample t-test

data : x and y

t = -0.36247 , df = 13.837, p-value = 0.7225

alternative hypothesis : True difference in mean is not equal to 0.

95 percent confidence interval :

-6.192719   3.692719

sample estimates :

mean of x   mean of y

12.125   12.875

Since, p-value > 0.05, we accept the H_0 at 5% level of significance

51

Q.4 The marks before & after a training program are given below:

before : 20, 25, 32, 28, 27, 36, 35, 25

After : 30, 35, 32, 37, 37, 40, 40, 23

Test the hypothesis, the training program is effective or not.

→ H_0 : Training program is not effective.

x = c (20, 25, 32, 28, 27, 36, 35, 25)

y = c (30, 35, 32, 37, 37, 40, 40, 23)

t.test (x, y, paired = t, alternative = "greater").

Paired t-test

data : x and y

t = -3.3859 , df = 7, p-value = 0.9942.

alternative hypothesis : true difference in mean is greater than 0.

95 percent confidence interval :

-8.467399 Inf

sample estimated :

mean of the differences

-5.75

Since, p-value > 0.05, we accept the H_0 at 5% level of significance

52

Q.5 Two random samples were drawn from two normal population & the values are:

A : 66, 67, 75, 76, 82, 84, 88, 90, 92

B : 64, 66, 74, 78, 82, 85, 87, 92, 93, 95, 97

Test whether the population have same variance at 5% level of significance.

→ H_0 : The variances of the two population are equal.

x = c (66, 67, 75, 76, 82, 84, 88, 90, 92)

y = c (64, 66, 74, 78, 82, 85, 87, 92, 93, 95, 97)

var.test (x, y)

F test to compare two variables

data : x and y

F = 0.70686, num df = 8, denom df = 10, p-value = 0.6359

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

95 percent confidence interval:

0.1833662 3.0360393

sample estimates :

ratio of variances

0.7068567

Since, p-value > 0.05, we accept the H_0 at 5% level of significance.

Scanned with CamScanner

Q6 The arithmatic mean of sample of 100 observations is 52. If the S.D is 7, test the hypothesis that the population mean is 55 or not at 5% level of significance.

→ H_0: population mean = 55

n = 100

m_x = 52

m_0 = 55

sd = 7

zcal = (mx - m0) / (sd / sqrt (n))

zcal

[1] -4.285714

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

pvalue

[1] 1.82153e-05

Since, p-value < 0.05, we reject the H_0 at 5% level of significance.

AM
12.2.20

Since, P-value < 0.05, we reject the H_0 at 5% level of significance.

Q2 Table below shows the relation between the performance of Mathematics & computer of cs students.

		Maths		
		HG	MG	LG
computer	HG	56	71	12
	MG	47	163	38
	LG	14	42	85

→ H_0: Performance between maths & computer is independent.

x = c(56, 47, 14, 71, 163, 42, 12, 38, 85)
m = 3
n = 3
y = matrix(x, nrow = m, ncol = n)
y

	[,1]	[,2]	[,3]
[1,]	56	71	12
[2,]	47	163	38
[3,]	14	42	85

pv = chisq.test(y)
pv

Pearson's chi-squared test

data: y

X-squared ... 145.78, df = 4, p-value ... 2.2e-16

Since, P-value < 0.05, we reject the H_0 at 5% level of significance.

Q3 perform ANOVA for the following data

Varieties	Observations
A	50, 52
B	53, 55, 53
C	60, 58, 57, 56
D	52, 54, 54, 55

→ H_0: The means of variety of A, B, C, D are equal

x_1 = c(50, 52)
x_2 = c(53, 55, 53)
x_3 = c(60, 58, 57, 56)
x_4 = c(52, 54, 54, 55)
d = stack(list(b1 = x1, b2 = x2, b3 = x3, b4 = x4))
names(d)
[1] "values" "ind"
oneway.test(values ~ ind, data = d, var.equal = T)

One-way analysis of means

data: values and ind.
F = 11.735, num df = 3, denom df = 9, P-value = 0.00183

anova = aov (values ~ ind, data = d)
summary (anova)

	DF	Sum Sq	Mean Sq	F values	Pr(>F)
ind	3	71.06	23.688	11.73	0.00183 **
Residuals	9	18.17	2.019		

--

Since, p-value < 0.05, we reject the H_0 at 5% level of significance.

Q4 Perform Anova for the following data:

Types	Observations
A	6, 7, 8
B	4, 6, 5
C	8, 6, 10
D	6, 9, 9

→ H_0: The means of types of A, B, C, D are equal.

x1 = c(6, 7, 8).
x2 = c(4, 6, 5).
x3 = c(8, 6, 10).
x4 = c(6, 9, 9).
d = stack (list (b1 = x1, b2 = x2, b3 = x3, b4 = x4))
names (d).
[1] "values" "ind"
oneway.test(values ~ ind, data = d, var.equal = T)

Practical No. 10

Topic :- Non. parametric test

Q. Following are the amount of sulphur oxide emitted by a factory.
17, 15, 20, 29, 19, 18, 22, 25, 27, 9,
24, 20, 17, 6, 24, 14, 15, 23, 24, 26.

Apply sign test, to test the hypothesis that the population median is 21.5
against the alternative that it is less than 21.5

→ H_0 : population median is 21.5

H_1 : It is less than 21.5

```
x = c(17, 15, 20, 29, 19, 18, 22, 25, 27, 9, 24, 20, 17,
     6, 24, 14, 15, 23, 24, 26)
m = 21.5
sp = length(x[x>m])
sn = length(x[x<m])
n = sp + sn
n
[1] 20
pv = pbinom(sp, n, 0.5)
pv
[1] 0.4119015
```

Since, p-value > 0.05, we accept the H_0 at 5% level of significance

Q.3 For the following data : 60, 65, 63, 89, 61, 71, 58, 51, 48, 66.
Test the hypothesis using wilcoxon's ~~test~~ signed ~~rank test~~ rank test. For testing the hypothesis that the median is 60 against the alternative it is greater than 60.

→ H_0: Median is 60.
H_1: It is greater than 60.

```
x = c(60,65,63,89,61,71,58,51,48,66)
mu = 60
wilcox.test (x, alter = "greater", mu = 60)
```

wilcoxon signed rank test with continuity correction

data: x
V = 29, P-value = 0.2386
alternative hypothesis : true location is greater than 60.

Since, P-value > 0.05 , we accept the H_0 at 5% level of significance.

[NOTE: If the alternative is less, we have to write ~~[illegible]~~ wilcox.test (x, alter = "less", mu = 60)
If the alternative is not equal to, we have to write wilcox.test (x, alter = "two.sided", mu = 60)]

Q.4 Test the hypothesis ; median is 12, against the alternative that it is less than 12. using wilcoxon's values : 12, 13, 10, 20, 15, 5, 1, 7, 6, 11, 9, 20.

→ H_0: Median is 12.
H_1: It is less than 12.

```
x = c(12,13,10,20,15,5,1,7,6,11,9,20)
mu = 12
wilcox.test (x, alter = "less", mu = 12)
```

wilcoxon signed rank test with continuity correction

data: x
V = 25, P-value = 0.2521
alternative hypothesis : true location is less than 12.

Since, P-value > 0.05 , we accept the H_0 at 5% level of significance.

An
4.3.20