Practical 1

Title: Random Variable

1] Find the mean & variance for the following

a]

x	-1	0	1	2
$P(x)$	0.1	0.2	0.3	0.4

Solution

x	$P(x)$	$x \cdot P(x)$	$E(x)^2$	$[E(x)]^2$
-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	~~0.8~~ 1.6	0.64
TOTAL	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 2$	$\Sigma = 0.74$

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

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

$= 2 - 0.74$

$= 1.24$

Mean $E(x) = 1$ & variance $V(x) = 1.24$

Scanned with CamScanner

b]

x	-1	0	1	2
$P(x)$	$1/8$	$1/8$	$1/4$	$1/2$

~~Solution~~

x	$P(x)$	$x \cdot P(x)$	$E(x)^2$	$[E(x)]^2$
-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	$\Sigma = 1$	$\Sigma = 9/8$	$\Sigma = 19/8$	$\Sigma = 69/64$

$\therefore$ Mean $= E(x) = \Sigma\, x \cdot P(x) = 9/8$

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

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

$$= \frac{152-69}{64}$$

$$= \frac{83}{64}$$

$\therefore$ Mean $E(x) = 9/8$ & Variance $V(x) = 83/64$

Scanned with CamScanner

Scanned with CamScanner

2]

x	-3	10	15
$P(x)$	0.4	0.35	0.25

Solution

x	$P(x)$	$x \cdot P(x)$	$E(x)^2$	$[E(x)]^2$
-3	0.4	-1.2	3.6	1.44
10	0.35	3.5	35	12.25
15	0.25	3.75	56.25	14.0625
TOTAL	$\Sigma = 1$	$\Sigma = 6.05$	$\Sigma = 94.85$	$\Sigma = 27.7525$

Mean $= E(x) = \Sigma\, x \cdot P(x) = 6.05$

Variance $= V(x) = \Sigma\, E(x)^2 - \Sigma[E(x)]^2$

$= 94.85 - 27.7525$

$= 67.0975$

$\therefore$ Mean $E(x) = 6.05$ & variance $V(x) = 67.0975$

a] 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

Solution As $P(x_i)$ is a pmf it should satisfy the properties of pmf which are

a] $P(x_i) > 0$ for all sample space

b] $\Sigma\, P(x_i) = 1$

Scanned with CamScanner

x	-1	0	1	2
$P(x)$	$k+1/13$	$5/13$	$1/13$	$k-4/13$

$$\therefore \sum P(x_i) = 1 = \frac{k+1}{13} + \frac{15}{13} + \frac{1}{13} + \frac{k-4}{13}$$

$$1 = \frac{k+1+15+1+k-4}{13}$$

$13 = 3k-2$

$15 = 3k$

$\boxed{k=5}$

x	$P(x)$	$xP(x)$	$E(x)^2$	$[E(x)]^2$
-1	$6/13$	$-6/13$	$6/13$	$36/\cancel{1424}\,169$
0	$5/13$	0	0	0
1	$1/13$	$1/13$	$1/13$	$1/\cancel{89}\,169$
2	$1/13$	$2/13$	$4/13$	$4/\cancel{129}\,169$
Total	$\Sigma = 1$	$\Sigma = -3/13$	$\Sigma = 11/13$	$\Sigma = 41/\cancel{129}\,169$

$\therefore$ Mean $= E(x) = \Sigma\, xP(x) = -3/13$

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

$= 11/13 - 41/\cancel{129}\,169$

$= 143-41/169$

$= 102/169$

$\therefore$ Mean $= -3/13$ & Variance $= 102/169$

The pmf of random variable x is given by

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

Obtain cdf. Find ① $P(-1 \le x \le 2)$ ② $P(1 \le x \le 5)$
③ $P(x \le 2)$ ④ $P(x \ge 0)$

Solution

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 $P(-1 \le x \le 2) = P(x \le 2) - P(x \le -1) + P(x = -1)$
$= F(x_b) - F(x_a) + P(a)$
$= F(2) - F(-1) + P(-1)$
$= 0.75 - 0.3 + 0.2$
$= 0.25$

② $P(1 \le x \le 5) = F(x_b) - F(x_a) + P(a)$
$= F(5) - F(1) + P(1)$
$= 0.95 - 0.65 + 0.2$
$= 0.15$

③ $P(x \le 2) = P(x = -3) + P(x = -1) + P(x = 0) + P(x = 1) + P(x = 2)$
$= 0.1 + 0.2 + 0.15 + 0.2 + 0.1$
$= 0.75$

④ $P(x \ge 0) = 1 - F(0) + P(0)$
$= 1 - 0.45 + 0.15$
$= 0.40$

Q.3] Let f be continuous random variable with pdf

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

Obtain cdf of x. Find mean

<u>Solution</u> By definition of cdf, we have

$$F(x) = \int_{-1}^{x} t\,dt$$

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

$$= \frac{1}{2}\left(\frac{1}{2}x^2 + x\right) \quad \text{for } -1 \le x \le 1$$

Hence the cdf is

$$F(x) = 0 \quad \text{for } x \le -1$$
$$= \frac{1}{4}x^2 + \frac{1}{2}x \quad \text{for } -1 \le x \le 1$$
$$= 0 \quad \text{for } x \ge 1$$

Practical 2

Title: Binomial Distribution

1] An unbiased coin is tossed 4 times. Calculate the probability of obtaining no head, atleast one head & more than one tail

NO HEAD:

```
> dbinom(0,4,0.5)
[1] 0.0625
```

ATLEAST ONE HEAD:

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

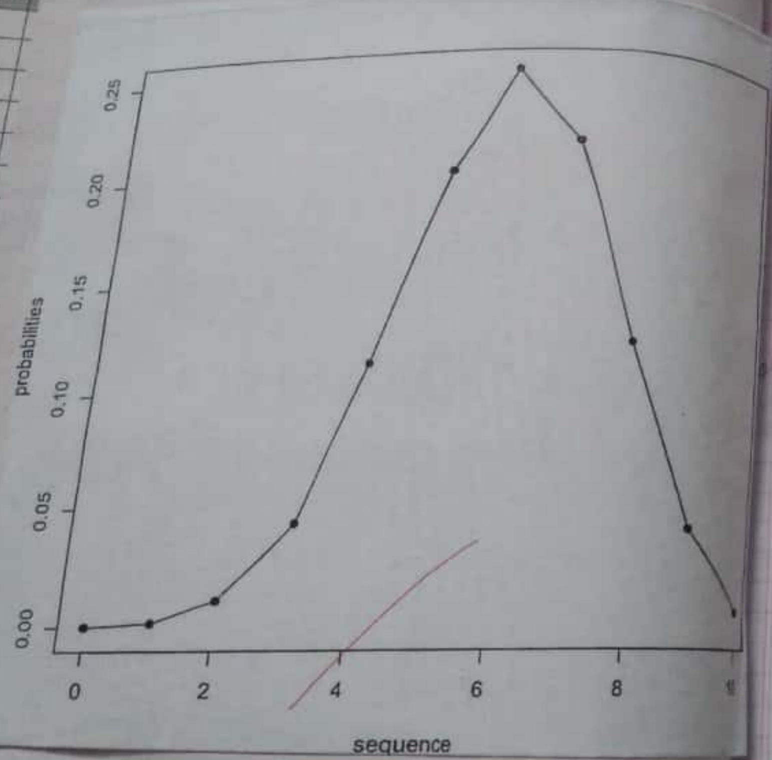
MORE THAN ONE TAIL:

```
> pbinom(1,4,0.5, lower.tail = F)
[1] 0.9375
```

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

```
> pbinom(2,5,0.3)
[1] 0.83692
```


Scanned with CamScanner



3] 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)
[1] 0.324135
> dbinom(3,6,0.3)
[1] 0.18522
> dbinom(2,6,0.3) + dbinom(3,6,0.3)+dbinom(4,6,0.3)
[1] 0.74373
```

4] For n=10, p=0.6, evaluate binomial probabilities and plot the graph 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   0.2006581248
    0.2508226560   0.2149908480   0.120932352
    0.0403107840   0.0060466176

> plot(x,y,xlab='sequence',ylab="probabilities"," o ",pc
```

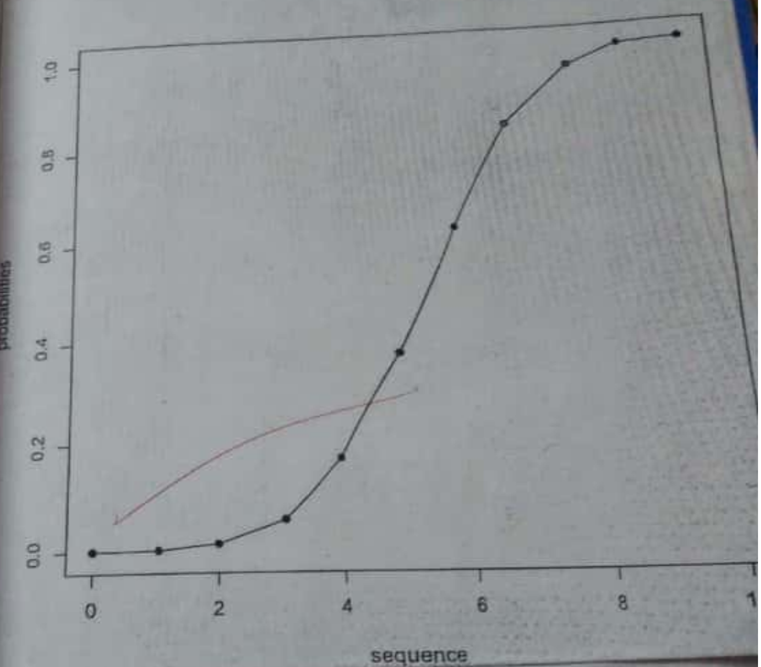
```
> x = seq (0,10)
> y = pbinom (x, 10, 0.6)
> plot (x, y, xlab = "sequence", ylab = "probabilities", "O", pch = 16)
```

5] Generate a random sample of size 10 for for a B D ~ B(1, 0.3) find the mean & the variance of the sample

```
> x = rbinom (8, 10, 0.3)
[1] 2 2 3 4 3 4 2 3
> Dbinom (8 summary (x)
[1] 2.7375
> var (x)
[1] 3.125
```

6] The probability of man hitting the target is 1/4 if he shoots 10 times. What is the probability that he hits the target exactly 3 times, probability that he hits target atleast one time

```
> dbinom (3, 10, 0.25)
[1] 0.2502823
> 1 - dbinom (1, 10, 0.25)
[1] 0.8122883
```



7) Bits are sent for communication channel in packet of 12. If the probability of bit being corrupted is 0.1. What is the probability of no more than 2 bits are corrupted in a packet?

> pbinom(2,12,0.1, lower.tail =F) + dbinom(2,12,0.1)
[1] 0.3400077

Practical 3

Topic: Normal Distribution

Q.1] A normal distribution of 100 students with mean 50 = 10

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)

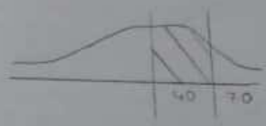
```
> pnorm(30,40,15)
[1]  0.2524925

> pnorm(70,40,15)- pnorm(40,40,15)
[1] 0.4772499

> pnorm(35,40,15) - pnorm(25,40,15)
[1] 0.2107861

> 1-pnorm(60,40,15)
[1] 0.0912122
```

#2

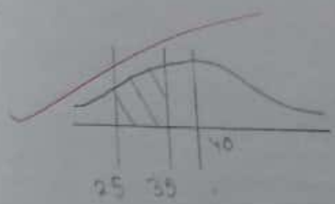
Q.1]

a]

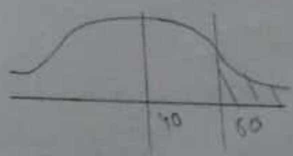
b]



c]



d]



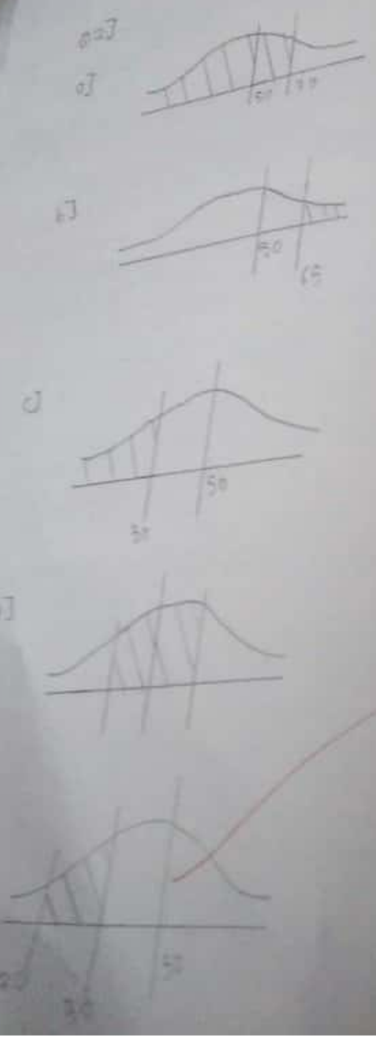
Q2]

a]

b]

c]

d]



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

```
> pnorm(70,50,10)
[1] 0.9772499

> 1- pnorm(65,50,10)
[1] 0.0668072

> pnorm(32,50,10)
[1] 0.03593032

> pnorm(60,50,10) - pnorm(35,50,10)
[1] 0.7745375

> pnorm(30,50,10) - pnorm(20,50,10)
[1] 0.02140023
```

2] Let X~N(160,400) find K1 & K2 such that P(X<K1)=0.6 P(X>K2)=0.8

```
> qnorm(0.6,160,20)
[1] 165.0669
> qnorm(0.8,160,20)
[1] 176.8324
```

Q.4] A random variable X follows normal distribution with $\mu = 0, \sigma = 2$. Generate 100 observation and evaluate its mean, median & variance

```
> x=rnorm(100,10,2)
> summary x
[1] Min    1st Q   Median   Mean    3rd Q    Max
    5.713  8.444   9.723    9.914   11.325   14.27
> var(x)
[1] 3.648924
```

Q.5] Write a command to generate 10 random numbers 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
```

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

```
> x = rnorm(100,10,2)
> summary x
[1] Min     1st Q   Median   Mean    3rd Q    Max
    5.713   8.444   9.723    9.914   11.328   14.73
> var(x)
[1] 3.648924
```

Q5] Write a command to generate 10 random numbers 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
```

Sample mean & deviation given single population

Q1) Suppose the food level on the cookie bag ... that it has almost 2gms of saturated salt in a single cookie. In a sample of 35 cookies, it was found that mean and of saturated salt per cookie is 2.1gm. Assume that its sample standard deviation is 0.3 at (5% level of significance) can be rejected the claim on food ... label.

$\to \sigma = 0.3$

$n = 35$

$\bar{x} = 2.1$

$x = 2$

$\to H_0$ (null Hypothesis) $= \mu \le 2$

$\to H_1$ (alt Hypothesis) $= \mu > 2$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\therefore z = \frac{2.1 - 2}{\frac{0.3}{\sqrt{35}}} = 1.972021$$

pvalue $= 1 - \text{pnorm}(z)$

$= 0.0243$

$\therefore$ Reject the null hypothesis $\because$ pvalue < 0.05

$\therefore$ Accepted alternate hypothesis

Scanned with CamScanner

Q2] A sample of 100 customers was randomly selected. It was found that average [illegible] was 275/-. The [illegible] using 0.05 level of significance, would you conclude that the amt spent by the customer is more than 250/- whereas the restaurant claim that it is not <250/-

⇒ $\bar{x} = 275$, $\mu = 250$, $\sigma = 30$, $n = 100$

$H_0 = \mu \le 250$

$H_1 = \mu > 250$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{275 - 250}{\frac{30}{\sqrt{100}}} = 8.333$$

⇒ pt(z, 99, lower tail = F)

pvalue = 2.3067 38e-13

⇒ Reject the null hypothesis ∵ pvalue < 0.05

Accept the alternate hypothesis ($\mu > 250$)

Scanned with CamScanner

Q3) A quality control engineer finds that sample of 100 have average life of 470 hours. Assuming population test whether the population mean is 480 hours vs population mean <480 hours at 1.05 → 0.05

$n = 100$, $\bar{x} = 470$, $\mu_1 \leq 480$, $\sigma = 25$, $\mu = 480$

→ $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = -4$

→ pt (z, 99, lower tail = F)

= 6.112576e-05

→ Reject the null hypothesis p<0.05

Accept the alternate hypothesis ($\mu_1 < 480$)

Q4) A principal at school claims that the IQ is 100 of the students. A random sample of 30 students whose IQ was found to be 112. The SD of population is 15. Test the claim of principal.

→ Method 1: 1 tail test

$x_0 = \mu = 100$

$x_1 = \mu = 100$

$\bar{x} = 112$, $SD = 15$, $\mu = 100$, $n = 30$

→ $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

$= \dfrac{112 - 100}{\frac{15}{\sqrt{30}}} = 4.38118$

Scanned with CamScanner

pvalue = 6.28567e-06

→ Reject the null hypothesis = claims of principle ($\mu = 100$)

Method 2: 2 tail-test

$\mu_0 = \mu = 100$

$\mu_1 = \mu = 100$

pvalue = 2x(1-pnorm(abs(z))) = 1.177134e-05

→ Reject the null hypothesis   pvalue = 0.05

(*) Single population proportion

Q] 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% LOC.

$$\Rightarrow z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

$\therefore p_0 = 0.5$

$q_0 = 1 - p_0 = 0.5$

$p = \frac{21}{40} = 0.7$

$n = 40$

$$\therefore z = \frac{0.7 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{40}}}$$

Scanned with CamScanner

Scanned with CamScanner

$H_0 - \mu = 0.5$

$H_1 = \mu \neq 0.5$

>> pvalue = 2x (1- pnorm(abs(z)))
∴ pvalue = 0.01141209

∴ Reject the null hypothesis ∵ $p < 0.05$
Accept the alternate hypothesis

c] In an hospital 480 females & 520 males are born in a
To confirm that male and female are equal in

$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$ $p \to \frac{520}{1000} = 0.52$, $p_0 = 0.5$, $q_0 = 0.5$, $n = 1000$

$H_0 = [p = p_0]$

$H_1 = [p \neq p_0]$

>> z = (p-p0) (sqrt(p0q0/n))
>> z = 1.2645

pvalue = 2 x (1- pnorm(abs(z)))
∴ pvalue = 0.2060506

Reject the null hypothesis ∵ pvalue < 0.5
∴ Accept the alternate hypothesis i.e. $p \neq p_0$

Scanned with CamScanner

Q] In a big city, 325 men out of 600 men are found to be self employed. Conclusion is that maximum men in city are self employed.

$$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

$p = 325/600 = 0.541607$, $p_0 = 0.5$, $q_0 = 0.5$, $n = 600$

$H_0 = [p_0 = p]$
$H_1 = [p \neq p_0]$

>>> z = ((0.5416 - 0.5)/(sqrt(0.5 * 0.5/600)))
>>> z = 2.037975

>>> pvalue = 2x(1 - pnorm(abs(z)))
>>> pvalue = 0.04155239

Reject the null hypothesis ∵ pvalue < 0.5
Accept the null hypothesis i.e. $p \neq p_0$

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

$$z = \frac{p - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

, $p = 0.05$ (50/400), $p_0 = 0.2$, $q_0 = 0.8$, $n = 400$

Scanned with CamScanner

$H_0 = [p = 0.2]$
$H_1 = [p \neq 0.2]$

$\Rightarrow z = (0.125 - 0.2) / (\text{sqrt}(0.2 * 0.8) / 400))$
$\Rightarrow z = -3.75$
$\Rightarrow$ pvalue = 2 x (1 - pnorm (abs(z)))
p value = 0.00017683 346

Reject the null hypothesis if pvalue < 0.2
Accept the null hypothesis if p ≠ 0.2

Formula:

$$z = \sqrt{pq\left(\frac{1}{n} + \frac{1}{m}\right)} \quad \text{where } p = \frac{p_1 n + p_2 n}{n + m}$$

11) In an election campaign, a telephone of 800 registered voters shows favour 460 second polls opinion 520 of 100 registered voters favoured the condition of 0.5%. LOC (Level of Confidence), is there sufficient evidence that popularity has decreased

$n = 800$, $p_1 = 460/800 = 0.575$, $m = 1000$, $p_2 = 520/1000 = 0.52$
$p = (0.575 * 800 + 0.52 \times 1000) / (1800)$
$p = 0.54444$

$z = \text{sqrt}(0.54 * 0.46) + 1/8$
$z = 0.001121394$

Scanned with CamScanner

$H_0 = p = 0.544$
$H_1 = p < 0.544$

pvalue = (2*(1-pnorm(abs(z))))
pvalue = 0.4491053

Accept the null hypothesis ∴ pvalue > 0.5
Accept p = 0.5444

Q27 From a consignment A, 200 articles are drawn & 44 were found defectives from consignment B, 200 samples are drawn out of which 30 are defective. Test whether the proportion of defective items in 2 consignments are significantly different.

$H_0 = p_1 = p_2$
$H_1 = p_1 \neq p_2$

$p_1 = 44/200 = 0.22$
$n = 200 = m$
$p_2 = 30/200 = 0.15$

>>> $p = \frac{(p_1 n + p_2 m)}{n+m}$

>>> p = (0.22*200 + 0.15*200)/400

>>> p = 0.125

z = (sqrt(0.185 * 0.815 * (2/200))
= 0.003882976

Scanned with CamScanner

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

∴ pvalue = 0.99696 12
pvalue > p

∴ Accept the null hypothesis i.e. $p_1 = p_2$

Nil

Scanned with CamScanner

Practical 5

Chi-Square Test

Q1] Use the following data to test whether the attributes conditions of home & child are independent

		Condition of Homes	
		Clean	Dirty
Condition of a child	Clean	70	50
	fclean	80	20
	Dirty	35	45

H_0 = Both are independent , H_1 = Both are dependent

```
> x = c(70,80,35)
> y = c(50,20,45)
> z = data.frame(x,y)
> z
[,]  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 = 2.698e-06

∴ Reject the null hypothesis

∴ Both are dependent

Q5) A dice is tossed 120 times & following results are obtained

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

Test the hypothesis that dice is unbiased

∴ H_0 = dice is unbiased, H_1 = 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] 0.956659
```

∴ Accept the null hypothesis
∴ Dice is unbiased

Scanned with CamScanner

Q5] An IQ test was conducted & the students were observed before & after training the result are following

Before	After
110	120
120	118
123	125
132	136
125	121

Test whether there is change in the IQ after the training.

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

```
>a = c(120, 118, 125, 136, 121)
>b = c(110, 120, 123, 132, 125)
>z = sum((b-a)^2)/a)
>pchisq(z, df = length(b)-1)
[1] 0.1135950
```

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

Scanned with CamScanner

	Graduate	Undergraduate
Online fav	20	25
No fav	40	5

Is there any association between student's preference for type of education & matrix

$\therefore H_0$ = Independent , H_1 = Dependent

> x = c(20,40,25,5)
> x = matrix(x, nrow=2)
> chisq.test(x)
Pearson's chi-squared test with Yates' continuity correction
data: x
X-squared = ~~23.977~~ 18.03, df = 1, p-value = ~~0.0000779~~ 2.15e-05

$\therefore$ Reject null Hypothesis
$\therefore$ Both are dependent

Scanned with CamScanner

Q5) A die is tossed 180 times.

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

Test the hypothesis that die is unbiased.

H_0 = die is biased
H_1 = die is unbiased

```
> x = c(20,30,35,40,12,43)
> chisq.test(x)
    Chi squared test for given probabilities

data: x
X-squared = 23.933, df = 5, p-value = 0.000223
```

∴ Reject null hypothesis
Die is unbiased.

Mesh

Scanned with CamScanner

Titled T Test

Q. Let $x = 3366, 3337, 3361, 3410, 3316, 3357, 3348, 3356, 3376, 3382, 3377, 3355, 3408, 3401, 3398, 3424, 3383, 3374, 3384, 3374$

write the R command for following to test hypothesis

1. $H_0: \mu = 3400$, $H_1: \mu \neq 3400$
2. $H_0: \mu = 3400$, $H_1: \mu > 3400$
3. $H_0: \mu = 3400$; $H_1: \mu \leq 3400$

at 95% level of confidence ~~also~~ Also check at 97% level of confidence

1. $H_0: \mu = 3400$

$H_1: \mu \neq 3400$

```
> x = c( 3366, 3337, 3361, 3410, 3316, 3357, 3348, 3356, 3376, 3382, 3377,
  3355, 3408, 3401, 3398, 3424, 3383, 3374, 3384, 3374)

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

One sample Test

data : x

t = -4.4865 , df = 19, p-value = 0.0002527

alternative hypothesis : true mean is not equal to 3400.

95 percent confidence level :

3361.797   3386.103

Sample Estimate:

mean of x :

3373.95

∴ Reject H_1
Accept H_0

> t.test (x, mu=3400, alter="two sided", conf.level=0.9)
one sample test

data: x
t = -4.4865, df=19, p-value = 0.000252 alternative
hypothesis: true mean is not equal to 3400
3360.43 3387.57

Sample estimates:
Mean of x
3373.95

∴ Reject H_0
Accept H_1

(5) $H_0: \mu = 3400$
$H_1: \mu > 3400$
> t.test (x, mu=3400, alter="greater", conf.level=0.95)
one sample t-test

data: x
t = -4.4865, df=19, pvalue = 0.9999
alternative hypothesis: true mean is greater than 3400
3363.91 Inf
Sample estimates:
mean of x
3373.95 ∴ Accept H_0
> t.test(x, mu=3400, alter="greater", conf.level=0.97)
one sided t-test

Scanned with CamScanner

data: x
t = -4.4865 , df = 19, pvalue = 0.9999
alternative hypothesis: true mean is greater than 3400
3367.337 Inf
sample estimates:
mean of x
3373.95

Accept H0.

Below are the data of gain in weights on 2 different diets A & B

Diet A : 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.

$\therefore H_0 = a-b = 0$

$H_1 = a-b \neq 0$

```
> a=c(25,32,30,43,24,14,32,24,31,31,35,25)
> 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 t-test

data : a and b
t = -0.62757 , df = 11 , p-value = 0.5429
alternative hypothesis: true difference in means is not equal to 0

Scanned with CamScanner

95 percent confidence interval:
-14.267330 7.933997
Sample estimates:
mean of the differences
-3.166667

∴ Accept H_0
There is no difference in weights

Q3] 11 students gave the test after 1 month they again gave the test after the tutions, do the marks give evidence that students have benefited by coaching

E1- 23,20,19,21,18,20,18,17,23,16,19
E2= 24,19,22,18,20,22,20,20,23,20,17
Test at 99 level of confidence

⇒ >e1 = 23,20,19,21,18,20,18,17,23,16,19
>e2 = 24,19,22,18,20,~~23~~,22,20,20,23,20,17
∴ 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 = -1.4132, df = 10, p-value = 0.08441
alternative hypothesis : true differences in mean is less than 10
99 percent confidence interval:
-Inf 0.867333

Scanned with CamScanner

Sample estimates:
Mean of the differences:
-1

∴ Accept Ho

Two drugs for BP was given & data was collected
D1 = 0.7, -1.8, -0.2, -1.2, -0.1, 3.4, 3.7, 0.8, 0, 2

D2 1.9, 0.8, 1.1, 0.1, -0.1, 4.4, 5.5, 1.6, 4.6, 3.4

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

→ Ho : d1 = d2
H1 : d1 ≠ d2
> d1 = c(0.7, -1.6, -0.2, -1.2, -0.1, 3.4, 3.7, 0.8, 0, 2)
> 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
t = -4.0621, df = 9, p-value = 0.002833
alternative hypothesis: true difference in means is not equal to 0.
95 percent confidence interval:
Mean of the difference:
-1.58

∴ Reject Ho
Accept H1

Scanned with CamScanner

63

Q7] If there is difference in salaries for the same job in 2 different countries
CA: 53000, 49051, 41974, 44366, 40470, 36963
CB: 62490, 58850, 49495, 52263, 47674, 43552

⇒ ∴ H_0 : $S1 = S2$
H_1 = $S1 \neq S2$

```
> ca = c(53000, 49051, 41974, 44366, 40470, 36963)
> cb = c(62490, 58850, 49495, 52263, 47674, 43552)

> 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 difference in means is not
  equal to 0
95 percent confidence interval:
     -10404.821    -2792.846
sample estimates:
mean of the differences:
           -6598.833
```

∴ Reject H_0
∴ Accept H_1

Practical - 7

F Test

1) Life expentancy in 10 region of Indian in 1990 & 2000 are given below. Test whether the variance of the 2 times are same

1990 = 37, 39, 36, 43, 45, 44, 46, 49, 50, 51
2000 = 44, 45, 47, 43, 42, 47, 50, 41, 48, 52, 42, 57

$H_0 = \sigma_1^2 = \sigma_2^2$
$H_1 : \sigma_1^2 \neq \sigma_2^2$
s1 = c(37, 39, 36, 43, 45, 44, 46, 49, 50, 51)
s2 = c(44, 45, 47, 43, 42, 47, 50, 41, 48, 52, 42, 57)
var.test(s1, s2) p value = 0.9176
Accept Ho

2)

I	25, 28, 26, 22, 22, 29, 31, 31, 26, 31
II	30, 25, 31, 32, 23, 28, 36, 26, 31, 32, 33, 27, 31, 38, 27)

→ $H_0 : \sigma_1^2 = \sigma_2^2$
$H_1 : \sigma_1^2 \neq \sigma_2^2$
s1 = c(25, 28, 26, 22, 22, 29, 31, 31, 26, 31)
s2 = c(30, 25, 31, 32, 23, 28, 36, 26, 31, 32, 33, 27, 31, 38, 27)
p value = 0.5341
∴ Accept Ho

3) For the following data test the hypothesis
i) equality of 2 population mean
ii) equality of proportion variance
→ $H_0 = \sigma_1^2 = \sigma_2^2$
$H_1 = \sigma_1^2 \neq \sigma_2^2$

Scanned with CamScanner

x=c(175, 168, 165, 190, 181, 185, 175, 200)
y=c(180, 170, 153, 180, 179, 183, 189, 205)
var.test(x,y)
p value = 0.7759

Accept H_0
t.test(x,y)
p value = 0.8216
Accept H_0

Q4] The following are the prices of commodity in the sample of shops selected at random from different cities

city A = 74.10, 77.70, 75.35, 74, 73.80, 79.30, 75.30, 76.80, 77..., 76.40

city B = 70.80, 74.90, 76.20, 72.80, 78.10, 74.70, 69.80, 81.20

→ x=c(74.10, 77.70, 75.35, 74, 73.80, 79.30, 75.30, 76.80, 77..., 76.40)

y=c(70.80, 74.90, 76.20, 72.80, 78.10, 74.70, 69.80, 81.20)

→ Shapiro.test(x)
p value = 0.6559
∴ data is normal

→ shapiro.test(y)
p value = 0.9304
∴ data is normal

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$
F value = 0.04249
∴ variances are not equal
∴ Reject H_0
accept H_1

$H_0: \bar{x}_1 = \bar{x}_2$
$H_1: \bar{x}_1 \neq \bar{x}_2$
>t.test(x, var.equal=F)
p.value = 3.485e-16
>t.test(y, var.equal=F)
p.value = 1.46 e-10
∴ Accept H_0

Q) Prepare a csv file in excel. Import the file in R & apply the test to check the equality of variance of 2 datas.
Obs1. 10, 15, 17, 11, 16, 20
Obs2. 15, 14, 16, 11, 12, 19.

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$

```
> Data = read.csv(file.choose(), header=T)
> Data
Observation1    Observation2
     10              15
     15              14
     17              16
     11              11
     16              12
     20              19

> attach(Data)
> mean(observation1)
  14.83333

> var.test(observation1, observation2)
  p value = 0.5717.
```

∴ Accept Ho.

Klert

Scanned with CamScanner

Practical 8

Title: Non Parametric Test

1] The times of failures in hrs of 10 randomly selected a volt battery of a certain company is as follows

28.9, 15.2, 28.7, 72.5, 48.6, 52.4, 37.6, 49.5, 62.1, 54.5

Test the hypothesis that the population median is 63 against alternative is than 63 at 5% of level of significance

H_0: median = 63

H_1: median < 63

```
>x=c(28.9,15.2,28.7,72.5,48.6,52.4,37.6,49.5,62.1,54.5)
>sp= length(which(x>63))
>sn= length(which(x<63))
>sp
4
>sn
1
>n=sp+sn
>qbinom(0.05,n,0.5)
2
```

∴ qbinom < sn

Accept Ho

∴ median = 63

Scanned with CamScanner

Q2] The following data gives the 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, 47, 49, 47, 55, 59, 63, 53, 56, 67, 49, 60,
64, 53, 50, 48, 51, 52, 54.
Use the sign test to test whether the median wt of population is 50kg against alternative it is >50kg

$\therefore H_0$: median = 50
$\therefore H_1$: median > 50

```
x=c(46,49,57,64,46,67,54,48,69,61,57,54,50,48,65,
61,66,54,50,48,49,62,47,49,47,55,59,63,53,56,61,
49,60,64,53,50,48,51,52,54)
>sp=length(which(x>50))
>sp
>25
>sn=length(which(x<50))
>sn
12
>n=sp+sn
>qbinom(0.05,n,0.5)
14
```

$\therefore$ qbinom > sn
$\therefore$ Reject H_0

Q) The median age of tourists visiting a certain place is claim to be 41 yrs. A random sample of 20 tourists have the age 25, 29, 52, 48, 57, 39, 45, 36, 30, 49, 28, 39, 44, 63, 32, 65, 42. Use the sign test to check the claim.

H_0: median = 41

H_1: median ≠ 41

```
>x=c(25,29,52,48,57,39,45,36,30,49,28,39,44,63,32,65,42)
>sp=length(which(x>41))
>sp
 9
>sn=length(which(x<41))
>sn
 8
>n=sp+sn
>qbinom(0.05,n,0.5)
> 5
```

∴ qbinom < sn

Accept H_0

median = 41.

Scanned with CamScanner

Q4) The time in minute that a patient has to wait for consultation is recorded as follows.
15,17, 24, 25, 20,21, 72,28, 12,25, 24,26
Use wilcoxon sign test to check whether the median weight in time is more than 20 at 5% level of significance.

→ H0: median>20
H1: median≤20
x = c(15,17,24,25,20,21,72,28,12,25,24,26)
wilcoxon.test(x,alternative="less")
p value = 0.909
Accept H0.

Q5) The weight in kg of the person before & after the

weight before	weight after
65	72
75	82
75	72
67	66
72	73

Use wilcoxon test to check whether the weight of person increases after stopping the smoking use 5% level of significance is 5% level of confidence

H0 : increases after ~~the~~ stopping smoking
H1 : does not increase after stopping smoking.

x=c(65,75,75,62,72)
y=c(72,52,72,66,73)

z=x-y

z
wilcoxon test(z,mu=0)
p value = 0.1736
accept Ho

Scanned with CamScanner

Practical

[illegible] Test

Q] The following data gives the effect of 3 treatments
T1 = 2, 3, 7, 2, 6
T2 = 10, 8, 7, 5, 10
T3 = 10, 13, 14, 13, 15

Test the hypothesis that all treatments are equally effected

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

   > stack(z)
      values  ind
   1     2    t1
   2     3    t1
   3     7    t1
   4     2    t1
   5     6    t1
   6    10    t2
```


Scanned with CamScanner

7	8	t2
8	7	t2
9	5	t2
10	10	t2
11	10	t3
12	13	t3
13	14	t3
14	13	t3
15	15	t3

```
> l=stack(z)
> aov(values~ind,data=l)
Call:
   aov(formula = values~ind, data=l)

Terms:
                      ind         Residuals
Sum of Squares     203.3333       54.0000
Deg. of Freedom       2              12

Residual standard error: 2.12132
Estimated effects may be unbalanced
> oneway.test(values~ind,data=l)

data: values and ind
F= 21.537, num df = 2.0000, denom df= 7.9314, p-value = 0.0006287
```


Scanned with CamScanner

[illegible]] The following data give life of tyres of four brands
A: 20, 23, 18, 17, 23, 24
B: 19, 15, 17, 20, 16, 17
C: 21, 19, 22, 17, 20
D: 15, 14, 16, 18, 14, 16
Test the hypothesis that average life of all tyres is same

```
-> > a=c(20,23,18,17,23,24
> b=c(19,15,17,20,16,17
> c=c(21,19,22,17,20
> d=c(15,14,16,18,14,16
> z=list(a,b,c,d)
> z
[[1]]
[1] 20 23 18 17 23 24

[[2]]
[1] 19 15 17 20 16 17

[[3]]
[1] 21 19 22 17 20

[[4]]
[1] 15 14 16 18 14 16

> L=list(i=a,j=b,k=c,l=d)
> L
```


Scanned with CamScanner

```
$i
[1] 20  23  18  17  23  23

$j
[1] 19  15  17  20  16  17

$k
[1] 21  19  22  17  20

$l
[1] 15  14  16  18  14  16

> a = stack(l)
> a
   values  ind
1
2
3
4
5
6
7
8
9
10
11
12
```

```
> aov(value ~ ind, data=c)
> oneway.test(value ~ind, data=c),
```

p-value = ~~0.7442~~ 0.04058

H_0 is rejected

Q] 3 types of wax is applied for the protection of car no. of days of protection were noted Test whether these are equally effective

→ H_0: The wax are equally effective
H_1: The wax are not equally effective

```
> a= c(44,45,46,47,48,49)
> b= c(40,42,51,52,55)
> c=c(50,53,58,59)

r= list(a1=a,b1=b,c1=c)
f= stack(r)
> oneway.test(value~ind, data=f)
```

p-value= 0.03822

Reject H_0

Scanned with CamScanner

Q) An experiment was conducted on 3 people and the observation were noted. Test whether the hypothesis that the group has equal ~~reacted~~ results on their mean

H_0 - The result are equal
H_1: The result are unequal

```
n= c(23,26,51,48,58,37,29,44)
n20=c(22,27,29,39,46,46,49,65)
n60=c(59,66,38,49,56,60,56,62)

m= list(a=n,b=n20,c=n60)
t=stack(m)

oneway.test(values~ind,data(t))
```

P-value 0.1412

H_0 is rejected