Practical 1

Title : Random Variable

Q) Find the mean and 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.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	1.6	0.64
	$\Sigma=1$	$\Sigma=1$	$\Sigma=2.0$	$\Sigma[E(x)^2]=0.74$

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

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

$= 2 - 0.74$

$= \underline{\underline{1.26}}$

$\therefore$ Mean $E(X) = 1$ & variance is 1.24

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	$\cancel{8}1$
	$\Sigma = 1$	$\Sigma = 9/8$	$\Sigma = 19/8$	$\Sigma = 69/64$

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

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

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

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

$\therefore$ Mean $[E(X)] = 9/8$ Variance $[V(X)] = \frac{83}{64}$

c)

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
	$\Sigma = 1$	$\Sigma = 6.05$	$\Sigma = 94.85$	$\Sigma = 27.7525$

$\therefore$ Mean $= E(X) = \Sigma x \cdot P(X)$

$$= 6.05$$

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

$$= 94.85 - 27.7525$$

$$= 67.0975$$

$\therefore$ Mean $E(X) = 6.05$ & Variance $V(X) = 67.097$

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

Solution:- As $P(X_i)$ is pmf, then it should satisfy the properties of pmf

a) $P(X) > 0$ for all sample space

b) $\Sigma(P(X)) = 1$

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

$\therefore \Sigma P(x) = 1 = \frac{k+1}{13} + \frac{k}{13} + \frac{1}{13} + \frac{k-4}{13}$

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

$13 = 3k - 2$

$15 = 3k$

$\boxed{k=5}$

X	$P(x)$	$X \cdot P(x)$	$E(x)^2$	$[E(x)]^2$
-1	$6/13$	$-6/13$	$6/13$	$36/169$
0	$5/13$	0	0	0
1	$1/13$	$1/13$	$1/13$	$1/169$
2	$1/13$	$2/13$	$4/13$	$4/169$
	$\Sigma = 1$	$\Sigma = -3/13$	$\Sigma = 11/13$	$\Sigma = 41/169$

$\therefore$ Mean $= E(x) = \Sigma x \cdot P(x) = -3/13$

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

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

$= \frac{143-41}{169}$

$= \frac{102}{169}$

Q3) 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) $P(1 \le x \le 5) = P(X \le 5) - P(X \le 1) + P(X = 1)$
$= F(X_b) - F(X_a) + P(a)$
$= F(5) - F(1) + P(1)$
$= 0.95 - 0.65 + 0.2$
$= 0.5$

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

Q4) Let f be continuous random variable with pdf

$\because f(x) = \frac{x+1}{2}, \quad -1 < x < 1$

$= 0$, otherwise

Obtain cdf of x. ~~Find mean~~

Solution:- By definition of cdf

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

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

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

Hence cdf is

$F(X) = 0$ for $-1 < x < 1$

$= \frac{1}{4}x^2 + \frac{1}{2}x$

$= 0$

Q5) Let f be continuous random variable with pdf

$f(x) = \frac{x+2}{18}, \quad -2 \le x \le 4$

$= 0$, otherwise

$\therefore$ For c.d.f.

Solution:- As per given condition

$$\therefore F(X) = \int_{-2}^{4} f\,dt$$

$$= \int_{2}^{4} \frac{x+2}{18}\,dx$$

$$= \int_{2}^{4} \frac{1}{18}\left(\frac{1}{2}x^2 + 2x\right) \ldots \text{ for } -2 < x < 4$$

$\therefore$ Cdf is $F(X) = 0$ for $X < -2$

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

for $-2 < x < 4$

$= 0 \ldots$ for $X \ge 4$

Practical 2

Title : Binomial Distribution

Q.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
```

Q.2) The probability that student is accepted to a prestigious college is 0.3. If 5 students appy, what is the probability that atmost 2 are selected.

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

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. Cal. P(X=2), P(X=3), P(1<x<5)

```
> dbinom(2,6,0.3)
[1] 0.324155
> 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
```

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 0.010616832O
    0.0424673280 0.1114767360 0.2006581248
    0.2508226560 0.2149908480 0.1209323520
    0.0403107840 0.0060466176

> plot(x,y,xlab="sequence", ylab="probabilities",
  "o", pch=16)
```

35) Generate a random sample of size 10 for a B.D. → B(8, 0.3). Find the mean & variance of sample.

```
> x = rbinom(8,10,0.3)
[1] 2 2 3 4 3 4 2 3
> summary(x)
[1] 2.375
> var(x)
[1] 3.125
```

36) The probability of a 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
```

37) 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) + pbinom(2,12,0.1)

[1] 0.3409977
```

Practical No. 3

"Normal Distribution"

A) normal distribution of 100 student with mean = 40,
SD = 15

Q) find no. of student whose marks are :- $P(x<30)$
$P(40<x<30)$
$P(25<x<35)$
$P(x>60)$

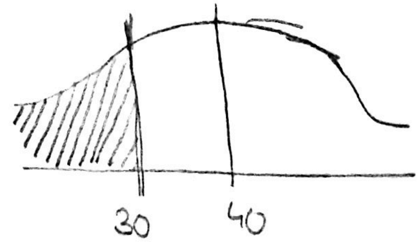
1) >pnorm (30, 40, 15)
[1] 0.25 24924

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

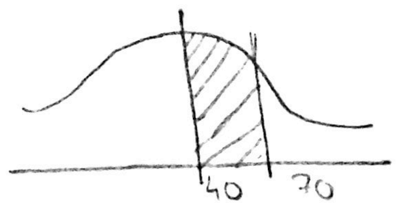
3) > pnorm (35, 40, 25) - pnorm (28, 40, 15)
[1] 0.2102861

4) 1 - pnorm (60, 40, 15)
[1] 0.09121192

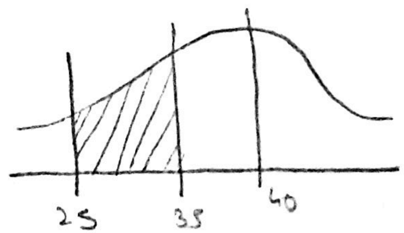
a)



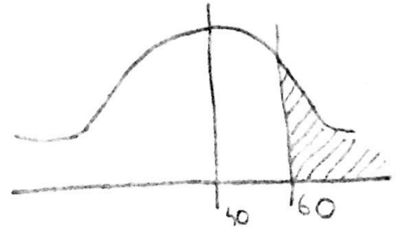
b)



c)

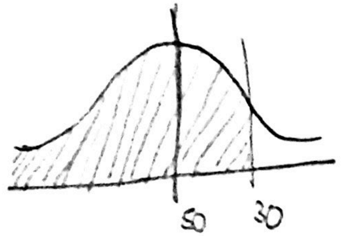


d)

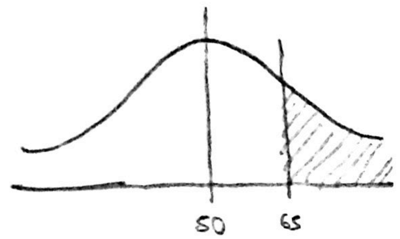


Q2

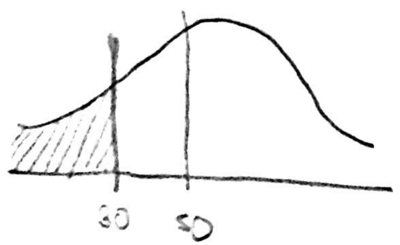
a)



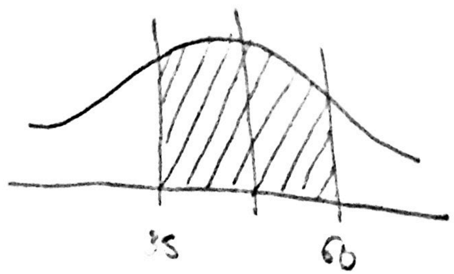
b)



c)



d)



Q2) If the random variable 'x' follows normal distribution with mean = 50, v = 100

Find:- i) $P(x \leq 70)$

ii) $P(x > 65)$

iii) $P(x \leq 32)$

iv) $P(35 < x < 60)$

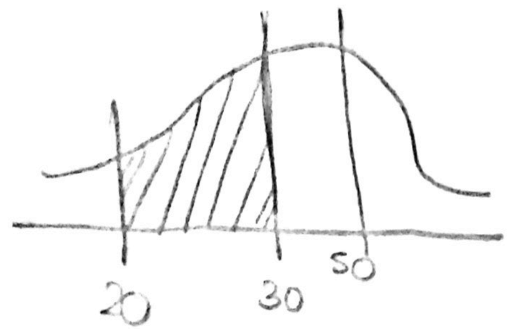
v) $P(20 < x < 30)$

```
> pnorm (70, 50, 10)
[1] 0.9712499

> 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
```

Q3) Dist nun (160, 400). Find k_1 & k_2 such that $P(X < k_1) = 0.6$ & $P(X > k_2) = 0.8$

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

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

Q2)
e)



Practical No. 4

Sample mean & standard deviation given single population.

Q1) Suppose the food level on the cookie bag states that it has almost 2 gms of saturated fat in a single cookie. In a sample of 35 cookies, it was found that mean ~~and~~ of saturated fat per cookie is 2.1gms Assume that the sample standard division deviation is 0.3 at 15% ~~of~~ level of significance can be rejected the claim on food label

↳ To check whether 'reject' or 'accept', null hypothesis at 95% level of confidence or 5% level of significance.

$\sigma = 0.3$

$n = 35$

$\bar{x} = 2.1$

$\mu = 2$

H_0 (null hypothesis) $= \mu < 2$

H_1 (alternate hypothesis) $= \mu > 2$

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

$$= \frac{21 - 2}{\frac{0.3}{\sqrt{35}}}$$

$$= 1.972027$$

∴ pvalue = 1 - pnorm (2)
= 0.243

∴ reject the null hypothesis
pvalue < 0.05
Accepted alternate hypothesis

Q2) A sample of 100 customers was randomly elected & it was found that average spending was 275%. The SD 30 using 0.05 level of significance, would conclude that the amount spend by the customer is more than 250/- whereas the restaurant claims that it is not 250/-

Ans: $\bar{x} = 275$, $u = 250$, $\sigma = 30$, $n = 100$

$H_0 = u < 250$

$H_1 = u > 250$

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

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

$= 8.333$

pnorm (2, 99, lower.tail = F)

∴ pvalue = 1.30576e-13

Q3) A quality control engineer finds a sample of 100 bulbs, having average life of 470 hours. Assuming population test whether the population mean < 480 hrs at LOS → 0.05

→ $n = 100$ $\bar{x} = 470$ $u < 480$ $\sigma = 25$

$$z = \frac{\bar{x} - u}{\frac{\sigma}{\sqrt{10}}}$$

pt (2, 99, lower.tail = T)

= 6.112576e.05

→ Reject the null hypothesis

∴ pt < 0.05

∴ Adapt the alternate hypothesis ($H_1 < 480$)

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

→ Method I:

$H_0 = u = 100$

$H_1 = u = 100$

$\bar{x} = 112$, S.D. = 15, $\mu = 100$, $n = 30$

$$z = \frac{\bar{x} - u}{\sigma / \sqrt{n}} = \frac{112 - 100}{15/\sqrt{30}}$$

$= 4.38178$

$\therefore$ pvalue = 5.88567e.06

$\therefore$ Project the null hypothsis = claim of principal

$[u = 100]$

→ Method 2:

$H_0 = M = 100$

$H_1 = u \neq 100$

$\therefore$ Pvalue = 2(1 - pnorm(abs(21)))

= 1.177134e.05

Reject the null hypothesis

$\therefore$ pvalue < 0.05

* Single Population Proportion:-

1) It is believed that coin is fair. The coin is tossed 40 times, 25 times head. Indicate whether the coin is fair or not at 95% L.O.C.

↳ Probability of sample

$$z = \frac{P - P_0}{\sqrt{\frac{p_0 q_0}{u}}}$$ - probability of population

$P_0 = 0.5$, $q_0 = 1 - P_0 = 0.5$

$P = \frac{28}{40} = 0.7$, $u = 40$

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

$\therefore u_0 = u = 0.5$

$\therefore u_1 = u \neq 0.5$

$\therefore$ Pvalue = 2(1 - pnorm(abs(2)))

= 0.01141209

Reject the null hypothesis

$\therefore p < 0.05$

$\therefore$ Accept the alternate hypothesis

Q) In an hospital 480 females & 520 males are born in a week to confirm male & ~~formate~~ female are equal in use.

$$z = \frac{P - P_0}{\sqrt{\frac{p_0 q_0}{n}}} \Rightarrow P = \frac{520}{1000} = 0.52 \quad P_0 = 0.5$$

$q_0 = 0.5$, $u = 1000$

$h_0 = [P = P_0]$

$h_1 = [P \neq P_0]$

$z = (P - P_0)$ (sqrt $(p_0 q_0 / u)$)

$z = 1.2645$

Pvalue = 2x(1 - pnorm(abs(2)))

= 0.2060506

$\therefore$ Reject the null hypothesis, < 0.5

$\therefore$ Accept the alternate hypothesis is $P \neq 100$

(Q2) In a big city 325 users of 600 users are found to be self employed. Conclusion is that maximum users in city are self employed

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

$p \rightarrow \frac{325}{600} = 0.541607$ $\quad p_0 = 0.5$

$q_0 = 0.5, \quad n = 600$

$h_0 = [p_0 = p]$
$h_1 = [p_0 \neq p]$

$z = [0.541607 - 0.5] / sqrt((0.5 * 0.5)/600)$

$z = 2.03975$

pvalue $= 2 \times (1 - pnorm(abs(z)))$
$= 0.04155239$

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

(Q3) 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 the hypothesis that experience of 10% of manufactured products.

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

$p = 0.125$ (50/400)
$p_0 = 0.2$ $\quad q_0 = 0.8$
$n = 400$

$\Rightarrow h_0 = [p = 0.2]$
$h_1 = [p \neq 0.2]$

$z = [(0.125 - 0.2) / sqrt((0.2 * 0.8)/400)]$
$= 0.00017 68 346$

Reject the null hypothesis: pvalue < 0.5
Accept the alternate hypothesis pvalue ≠ 0.2

* Formula:-

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

(Q1) In an election campaign a telephone of 800 ~~required~~ registered were ~~stos~~ pvalue 460. Second ~~pole~~ poll opinion 620 of 1000 registered voters favored the candidate at 0.5% L.O.C is there sufficient evidence that probability has decreased

$n = 800, \; p_1 = 460/800 = 0.575$
$m = 1000, \; p_2 = 620/1000 = 0.52$

$p = (0.575 * 800 + 0.52 \times 1000)/1800$

$\boxed{P = 0.5444}$

$z = 0.01121594$

$H_0 = p = 0.544$

$H_1 = p < 0.544$

Pvalue $= 2 \times (1 - \text{pnorm}(\text{abs}(z)))$

Pvalue $= 0.9991053$

∴ Accept the null hypothesis

∵ Pvalue > 0.5

∴ Accept $P = 0.5444$

Q2) From a consignment A 100 article are drawn & 44 ~~was~~ found defective from consignment 200 samples are drawn out of which 50 are defective. Test whether the proportion of defective items in 2 consignments are significantly different.

→ $H_0 = P_1 = P_2$

$H_1 = P_1 \neq P_2$

$P_1 = 44/200 = 0.22$

$n = 200 = m$

→ $P_2 = 30/200 = 0.15$

$p = \dfrac{(P_1 n + P_2 m)}{n+m}$

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

$p = 0.125$

$z = (\text{sqrt}(0.185 * 0.815) * (2/200))$

$= 0.00388976$

∴ pvalue $= 0.9969018$

∵ Pvalue $> p$

∴ Accept the null hypothesis i.e $P_1 = P_2$

Practical 5

Title : Chi-square Test

Q1) A Use the following data to test whether the attribute condition of home & child are independent.

Condition of Homes

Clean	Dirt
70	30
80	20
35	45

Ho = Both are independent , H_1 = Both are dependent

```
> x=c( 70, 80, 35)
> y=c( 50, 20, 45)
> z=data.frame(x,y)
> z
[1]  x    y
     70   30
     80   20
     35   45
```

```
> chisq.test(z)
```

∴ Reject null hypothesis

∴ Both are dependent

Q2) A dice is tossed 120 times & the following result are obtained

No. of items	Freq
1	30
2	25
3	18
4	10
5	22
6	15

Test the hypothesis that the dice is unbiased.

∴ Ho = dice is unbiased

H_1 = dice is biased

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

```
z = sum((obs) -exp)^2 /(exp))
Pchisq (Z, df =length(obs) -1)
```

∴ Accept the null hypothesis

∴ dice is unbiased.

Q3) A IQ test was conducted & the students were observed before & after training, the results are following

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

Test whether there is a change in IQ after the training

∴ H_0 = No change in IQ

∴ H_1 = Change in IQ

```
a = c(120, 118, 125, 136, 121)
b = c(110, 120, 123, 132, 125)
z = sum((a-b)^2)/a)
> pchisq(z, df = length(b)-1)
[1] 0.7135959
```

∴ There is change in IQ after training.

Q4)

	Graduate	Undergraduate (U.G.)
Online	20	25
Before	40	5

Is there any association between student ~~interface~~ preference for type of education & method:-

∴ H_0 : Independent

H_1 : Dependent

```
X = c(20, 40, 25, 5)
z = matrix(X, nrow = 2)
> chisq.test(z)
```

data: z

x-squared = 18.5, df = 1, pvalue = 2.157e-05

∴ Reject null hypothesis

∴ Both are independent.

F.P

Q5) A dice is tossed 180 times.

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

Test the hypothesis that dice is unbiased

$H_0 \to$ dice is biased
$H_1 \Rightarrow$ dice is unbiased

```
> x = c(20, 30, 35, 40, 12, 43)
chi.test(x)
data: x
```

X-squared = 23.933, df = 5, p-value = 0.000.

∴ Reject null hypothesis

∴ Dice is unbiased.

Practical 6: t-test

Title: t-test

1) 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 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 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, 3423, 3383, 3374, 3384, 3374
> t.test(x, mu = 3400, alter = "two sided",
conf.level = 0.95)
```

One sample t-test

data: x
t = -4.4865, df = 19, p-value = 0.0002528
alternative hypothesis: true mean it is not equal to 3400
95 percent confidence level: 3361.797 3386.103

Sample estimates:
mean of x: 3373.95
∴ Reject Ho
∴ Accept H_1
> t.test (x, mu = 3400, alter = "two.sided",
conf. level = 0.97)
One sample t-test

data: x
t = -4.4865, df = 19, p-value = 0.000252
alternative hypothesis: true mean is not equal
to 3400
97 percent confidence level: 3360.33 3387.57
Sample estimates
Mean of x: 3373.95
∴ Reject Ho
∴ Accept H_1

(2) $H_0 : \mu = 3400$
$H_1 : \mu > 3400$
> t.test (x, mu = 3400, alter = "greater", conf.
level = 0.95)
One sample t-test
data: x
t = -4.4865, df = 19, p-value = 0.9999
alternative hypothesis: true mean is greater than
3400
3363.91
Sample estimates
Mean of x: 3373.95
∴ Accept Ho

> t.test (x, mu= 3400, alter= "greater", conf.level=0.97)

One sided t-test

data: x

t= -4.4865, df= 19, p-value =0.9999

alternative hypothesis: true mean is greater than 3400

3367.337 Inf

Sample estimates:

Mean of x: 3373.95

∴ Accept H_0

(3) $H_0 : M = 3400$

$H_1 : M < 3400$

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

One sided t-test

data: x

t= -4.4865, df=19, p-value= 0.0001265

alternative hypothesis: true mean is less than 3400

95 percent confidence level

-Inf 3383.99

Sample estimates.

Mean of x:

3373.95

∴ Reject H_0

∴ Accept H_1

> t.test(x, mu=3400, alter= "less", conf.level=0.97)

One sided t-test

data: x

t = -4.4865, df = 19, p-value = 0.0001264

alternative hypothesis: true mean is less than 3400

97 percent confidence level

-Inf 3385.563

Sample estimates:

Mean of x:

3373.95

∴ Reject H_0

∴ Accept H_1

Q2) Below are the data of gain in weights of 2 d different data A & B.

Data A: 25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25

Data 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 t-test

data: a and b

alternative hypothesis: true difference in mean is not equal to 0

95 percent confidence level

-14.267330 7.933997

Sample estimates:

Mean of the differences:

-3.166667

∴ Accept Ho

∴ There is no difference in weights

Q5) 11 students gave the test after 1 month they again gave the test after coaching, do the marks gives evidence 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, 17

test at 99% level of ~~confidade~~ confidence.

> E1 = c(23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19)

> E2 = c(24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 17)

∴ H_0: $E_1 = E_2$

∴ H_1: $E_1 < E_2$

> t.test(E1, E2, paired = T, alter = "less", conf.level = 0.99)

Paired t-test

data : E1 and E2
alternative hypothesis :
t = -1.4832, df = 10, p-value = 0.08441
alternative hypothesis: true ~~mean~~ difference in means less than 0
99 percent confidence level.
-Inf 0.863333
Sample estimates :
Mean of the differences:
-1

∴ Accept H0

Q4) Two drugs for BP was given and 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

$H_0 : D_1 = D_2$
$H_1 : D_1 \neq D_2$

```
>D1=c(0.7,-1.6,-0.2,-1.2,-0.1,3.4,3.7,0.8,0.2
>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 level
Mean of the difference:
-1.58
∴ Reject H0
∴ Accept H1

(5) If there is difference in salaries for the same job in 2 different counteries.

CA : 53000, 49958, 41974, 44366, 40470, 36963
CB : 62490, 58850, 49495, 52263, 47674, 43552

∴ H_0 : CA = CB
∴ H_1 : CA ≠ CB
```
> CA = c(53000, 49958, 41974, 44366, 40470, 36963)
> CB = c(62440, 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 level
-10404.821 -2792.846

Sample estimates:
Mean of the differences:
-6598.833

∴ Reject H_0
∴ Accept H_1

Practical 7 (f-test)

1) Life expectancy in 10 regions in India in 1990 & 2000 are given below:
Test whether the variance of the two times are same:

$H_0 : \sigma(y1) = \sigma(y2)$
$H_1 : \sigma(y1) \neq \sigma(y2)$

1990: 37, 39, 36, 42, 45, 44, 46, 49, 50, ~~51~~, 52)
2000: 44, 45, 47, 43, 42, 49, 50, 41, 52, 59)

→ y1 = c(37, 39, 36, 42, 45, 44, 46, 49, 50, 42)
→ y2 = c(44, 45, 47, 43, 42, 49, 50, 41, 52, 59)

→ var.test(y1, y2)

F test to compare two variances

data: y1 and y2
F = 3.4085, num df = 10, denom df = 10, ~~p-value~~
p-value = 0.06609
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence level:
0.9170663   12.6688513
Sample estimates:
ratio of variances
3.408545

∴ Accept H_1

Q2) For the foll. data of the two samples. Test whether the variance ~~set~~ of the two sample are same or

a: (25, 28, 26, 22, 22, 29, 31, 31, 26, 31)
b: (30, 25, 31, 32, 23, 25, 36, 25, 31, 32, 32, 27
31, 38, 24)

$H_0: \sigma(a) = \sigma(b)$
$H_1: \sigma(a) \neq \sigma(b)$

```
> a=c(25,28,26,22,22,29,31,31,26,31)
> b=c(30,25,31,32,23,25,36,25,31,32,32,27
    31,38,24)

> shapiro.test(a)
        Shapiro-Wilk normality test
data: a
W = 0.89253, p-value = 0.181

> shapiro.test(b)
        Shapiro-Wilk normality test
data: b
W = 0.92642, p-value = 0.2411

> var.test(a,b)
```

F test to compare two variances

data: a and b

F = 0.60558 , num df = 9 , denom df = 14, ~~p-value~~

p-value = 0.4535

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

95 percent confidence interval:

0.1886943 2.2999517

Sample estimates:

ratio of variances

0.6055767

∴ Accept H_1

Q) For the foll. data test the hypothesis for:

① Equality of 2 population mean

② Equality of 2 population variance

I: 175, 168, 148, 190, 181, 185, 175, 200

II: 180, 170, 153, 180, 179, 183, 187, 205

$H_0: \sigma(I) = \sigma(II)$ $H_0: m(I) = m(II)$

$H_1: \sigma(I) \neq \sigma(II)$ $H_1: m(I) \neq m(II)$

I=c(175,148,168,190,181,185,175,200)

II=c(180,170,153,180,179,183,187,205)

> shapiro.test (I)
Shapiro-Wilk normality test
data : I
W = 0.96943 , p-value = 0.8943

> shapiro.test (II)
Shapiro-Wilk normality test
data: II
W = 0.96112 , p-value = 0.8283

for mean
(1) > t.test (I, II)
Welch Two Sample t-test
data : I and II
t = -0.32965 , df = 11.979 , p-value = 0.74
alternative hypothesis : true difference in mea
is not equal to 0
95 percent confidence interval:
-21.74526 16.03098
Sample estimates.
Mean of x and y
176.7143 179.5714

∴ Reject Ho
∴ Accept H_2

(2) > var.test (I, II)

F test to compare two variances

data : I and II

F = 1.0871 , num df = 6, denom df = 6 , p-value = 0.9218

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

95 percent confidence interval :

0.1867998 6.3268267

Sample estimates :

Ratio of variances :

1.087129

2) The foll are prices of commodity in the sample of shops selected at random from 2 differnt cities

City A = 74.10, 77.70, 75.35, 74, 73.80, 79.30, 75.80, 76.80, 79.10, 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, 75.80, 76.80, 77.70, 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 ~~re~~ normal

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

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

P. value = 0.04249

2 varianas are not equal

∴ Reject Ho

∴ Accept H_1

$H_0 : \overline{X_1} = \overline{X_2}$

$H_1 : \overline{X_1} \neq \overline{X_2}$

> t.test (x, var.equal = F)

p value = 3.488e-16

> t.test (y, var. equal = F)

pvalue = 1.46 e-10

∴ Accept Ho

Q5) Prepare a .csv file in excel. Import the file in R and apply the test to check the equality of variancy of 2 data.

Obs1 : 10, 15, 17, 11, 16, 20 $\quad H_0 : \sigma_1^2 = \sigma_2^2$

Obs2 : 15, 14, 16, 11, 12, 19 $\quad H_1 : \sigma_1^2 \neq \sigma_2^2$

> data = read.csv (file.choose (), header = T)

> data

Obs..1	Obs..2
10	15
15	14
17	16
11	11
16	12
20	19

> attach(data)
> mean(obs..1)
14.83333

> var.test(obs..1, obs..2)
p-value = 0.5717

∴ Accept H_0

P.G

Practical No. 8

Non-Parametric Test

Q1) The times of failure (in hours) of 10 randomly selected 9V batteries of a certain company are 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 population median is less than 63.

Solution: 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)
> Sn = length(which(x<63))
> Sp = length(which(x>63))
> n = Sn + Sp
> qbinom(0.05, n, 0.5)
  2
> Sn
  9
#Since qbinom < Sn, Ho is accepted

> pbinom(Sn, n, 0.5)
  0.9990234
#Since pbinom > 0.5, Accept Ho
```

∴ Ho is accept

ii) Data of weights of 40 students is given. Test the hypothesis that the median is 50 against it is greater than 50

Solution:

H_0 : median = 50
H_1 : median > 50

```
> x=c(46,49,57,64,46,67,54,48,49,61,57,
  50,48,65,61,66,54,50,48,49,62,
  47,49,47,55,59,63,53,56,57,
  49,60,64,53,50,48,51,52,54)
> Sn = length(x[x<50])
> Sp = length(x[x>50])
> n = Sp + Sn
> qbinom(0.05,n,0.5)
  14
> Sn
  13
# Since qbinom > Sn, Reject Ho
```

∴ H_0 is rejected.

8) The median age of tourist visiting the certain place is to be 41 years. A random sample of 17 ~~por~~tourists is given. Use the sign test to check the claim.

Solution: Ho : ~~for~~ median = 41
H1 : median ≠ 41

```
> x=c(25,24,48,52,57,39,45,36,30,49,
     28,39,44,63,32,65,42)

> Sp=length(which(x>41))
> Sn=length(which(x<41))
> qbinom(0.05,n,0.5)
  5
> Sn
  8
#
# Since qbinom < Sn, Ho is accepted
```

∴ Median = 41
∴ Ho is accepted

5) The time in minutes that a patient have to wait for consultation is recorded. Use Wilcox'n ~~signed~~ test to test which if median is more than 20 at 5% level of significon

<u>Solution:</u> H_0 : median > 20
H_a : median < 20

> x=c(15, 17, 24, 25, 20, 21, 32, 28, 12, 25, 24, 26)

> Wilcox.test(x, alternative="less")
0.6999

∴ Accept H_0

5) The weight of person before and after quitting smoking is given. Check whether the weight increases after he stops smoking. Using Wilcox'n test at 5% level of significanc-

> b=c(65, 75, 75, 62, 72)
> a=c(72, 82, 72, 66, 74)

> x = b-a
> x
-7 -7 3 -4 -1

8.3

> Wilcox.test (x, mu = 0)

p-value = 0.1756

Since pvalue > 0.05

∴ Accept H_0

Practical 9 (ANOVA Table)

The foll. data gives the effect of three treatments

Treatment 1	Treatment 2	Treatment 3
2	10	10
3	8	13
7	7	14
2	5	13
6	10	15

```
t1=c(2,3,7,2,6)
t2=c(10,8,7,5,10)
t3=c(10,13,14,13,15)
```

H_0: Treatments are equally effective
H_1: Treatments are not equally effective

```
d=data.frame(t1,t2,t3)
e=stack(d)
oneway.test(values~ind, data=e)
```

p-value = 0.0006282

There is enough evidence to reject H_0.

Q2) The foll. is life of tyres of 4 brands

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

→ Test the hypothesis that the avg life all brands is same

Solution H0: Life of all brands of tyres is same
H1: Life of all brands of tyres is differ

```
> A=c(20, 13, 18, 17, 22, 24)
> B=c(14, 15, 17, 20, 16, 17)
> C=c(21, 19, 22, 17, 20)
> D=c(15, 14, 16, 18, 14, 16)
> Z=list(p=A, q=B, r=C, s=D)
> K=stack(Z)
> Oneway.test(values~ind, data=K)
```

p-value = 0.01762

∴ H0 is rejected

3) Three brands of wax applied for protection of car and no. of days ~~p~~ of protection were noted. Test whether these are equally effective

H_0: Equally Effective H_0: Not equally effective

> w1 = c(44, 45, 46, 37, 48, 49)

> w2 = c(40, 42, 51, 52, 55)

> w3 = c(50, 53, 58, 59)

> u = list(t = w1, f = w2, h = w3)

> i = stack(u)

> oneway.test(values ~ ind, data = i)

p-value = ~~0.3822~~ 0.03822

∴ We have enough evidence to reject H_0.

4) The experiment was conducted on 8 persons & the observation was noted

I : No exercise

II : 20 mins of exercise

III : 60 mins of exercise

Test the hypothesis that all groups have equal results on exercise

H_0: The results are equ

H_1 : The results are not equal

> I = c(23, 26, 51, 48, 58, 37, 29, 44)

> II = c(22, 27, 29, 39, 46, 48, 49, 65)

> III = c(59, 66, 38, 49, 56, 60, 56, 62)

> z1 = list(r1 = I, r2 = II, r3 = III)
> stk = stack(z1)
> oneway.test(values~ind, data = stk)

p-value = 0.01633

∴ Ho is rejected