Experiment No: 01

Experiment Name: Find out the point estimate of the population mean and interval estimation of the population mean. where 30 students quiz test marks is

(2, 4, 3, 23, 25, 27, 28, 13, 15, 16, 20, 14, 35, 33, 32, 21, 35, 40, 42, 22, 33, 13, 17, 20, 25, 29, 27, 40, 38, 31).

Total marks 50. Here population size N=30 and sample size n=10.

Also illustrate the sample size determination, sampling distribution for mean and check the unbiasedness of the population mean.

Objectives:

1. To calculate the point estimation and interval estimation.

2. To calculate sampling distribution for mean.

Scanned with CamScanner

3. To check the unbiasedness of the population mean.

4. To comment on the data.

Procedure:

Step-1: First of all we find out the population mean and population variance. Population length is N.

mean, $\bar{x} = \frac{\Sigma x_i}{N}$

variance, $s^2 = \frac{1}{N-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{N}\right]$

Step-2: To calculate point estimation and interval estimation.

interval estimation:

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \;,\; \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

Scanned with CamScanner

step-3: Sampling Distribution for mean. We choose the sample size $n=10$ from the population size $N=30$ Then we calculate the mean and unbiasedness.

bias = mean(nsample) - mean(population)

= 0

when bias is 0 then we can say the mean is unbiasedness.

step-4: Sampling Distribution for median. We choose the sample size $n=10$ from the population size $N=30$ The we calculate the median and unbiasedness.

bias = median(nsample) - median(population)

when bias is 0 then we can say the median is unbiasedness.

Scanned with CamScanner

step-5: Efficiency check

we calculate the mean and the median of sampling distribution.

Mean and median to be two unbiased estimator then which variance is ~~more~~ less than other then we say that this is more efficient than other.

R-Source code:

```
IQ<-c(2,4,3,23,25,27,28,13,15,16,20,14,
35,33,32,21,35,40,42,22,33,13,17,20,25,
29,27,40,38,31)
mean(IQ)
var(IQ)
length(IQ)

set.seed(1246)
x<-sample(IQ,10, replace=TRUE)
mean(x)
sd(IQ)
qnorm(0.025,0.1)
```

Scanned with CamScanner

```
## lower class interval
21.6-((1.96*11)/sqrt(10))

## upper class interval
21.6+((1.96*11)/sqrt(10))

##Sampling Distribution for mean
choose(30,10)
nsample<-req(0,300000)
for(i in 1:300000){
    nsample[i]<-(mean(sample(IQ,10,
                       replace=TRUE)))

}
mean(nsample)
bias=mean(nsample)-mean(IQ)

##Sampling Distribution for median
choose(30,10)
nsample2<-req(0,300000)
for(i in 1:300000){
    nsample[i]<-(median(sample(IQ,10,
                       replace=TRUE)))

}
```

Scanned with CamScanner

```
median(IQ)
median(nsample2)
bias = median(nsample2) - median(IQ)

### Efficiency check ###
L1 <- length(nsample)
V1 <- sum((nsample - mean(IQ))^2)/L1
V1
L2 <- length(nsample2)
V2 <- sum((nsample2 - median(IQ))^2)/L2
V2
```

Input and output:

```
mean(IQ) = 24.1
var(IQ) = 121.2655
length(IQ) = 30
mean(x) = 21.6
sd(IQ) = 11.012
qnorm = -1.96
14.78 # lower class interval
28.41 # upper class interval
```

Scanned with CamScanner

mean(nsample) = 24.097

bias = -0.0024

median(IQ) = 25

median(nsample2) = 25

bias = 0

L1 = 300000

V1 = 11.69

L2 = 300000

V2 = 19.97

Comment : From the R code we can see that the mean is a unbiased estimator and the median also unbiased estimator. The variance of ~~mean is~~ nsample is less then the variance of nsample2. So, the mean is more efficient ~~the~~ than median.

Experiment No: 02

Experiment Name: Two dice rolled, S is the sum of both faces. Find the expectation of S, E(S) and variance of S, V(S). Plot the distribution of S and dice D.

Objectives:

1. To find the expectation of S.
2. To find the variance of S.
3. To plot the distribution of S and dice D.
4. To comment on the data.

Procedure:

Step-1: Two dice rolled, S is the sum of both faces. To calculate the expectation of S, E(S).

Scanned with CamScanner

Step-2: To calculate the variance of S, V(S).

$$V(S) = \frac{\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right]}{n-1}$$

Step-3: To plot the distribution of S and dice D.

R-Source code:

```
S<- 2:12
A <-c(1:6,5:1)
PS<-C(1:6,5:1)/36
ES<-Sum (S*PS)
varS<-Sum((S-c(ES))^2*PS)

## plot distribution of S
barplot(PS, ylim=c(0,0.2),
        ylab="Probability",
        xlab="S",
        col="steelblue",
        Space=0,
        main="Sum of two dice rolls")
```

Scanned with CamScanner

```
#### plot distribution of D
Probability <- rep (1/6,6)
names (Probability ) <- 1:6
barplot (probability ,
         ylim = c(0,0.2),
         xlab = "D"
         col = "steelblue",
         Space = 0,
         main = "outcomes of a single
                          dice roll")
```

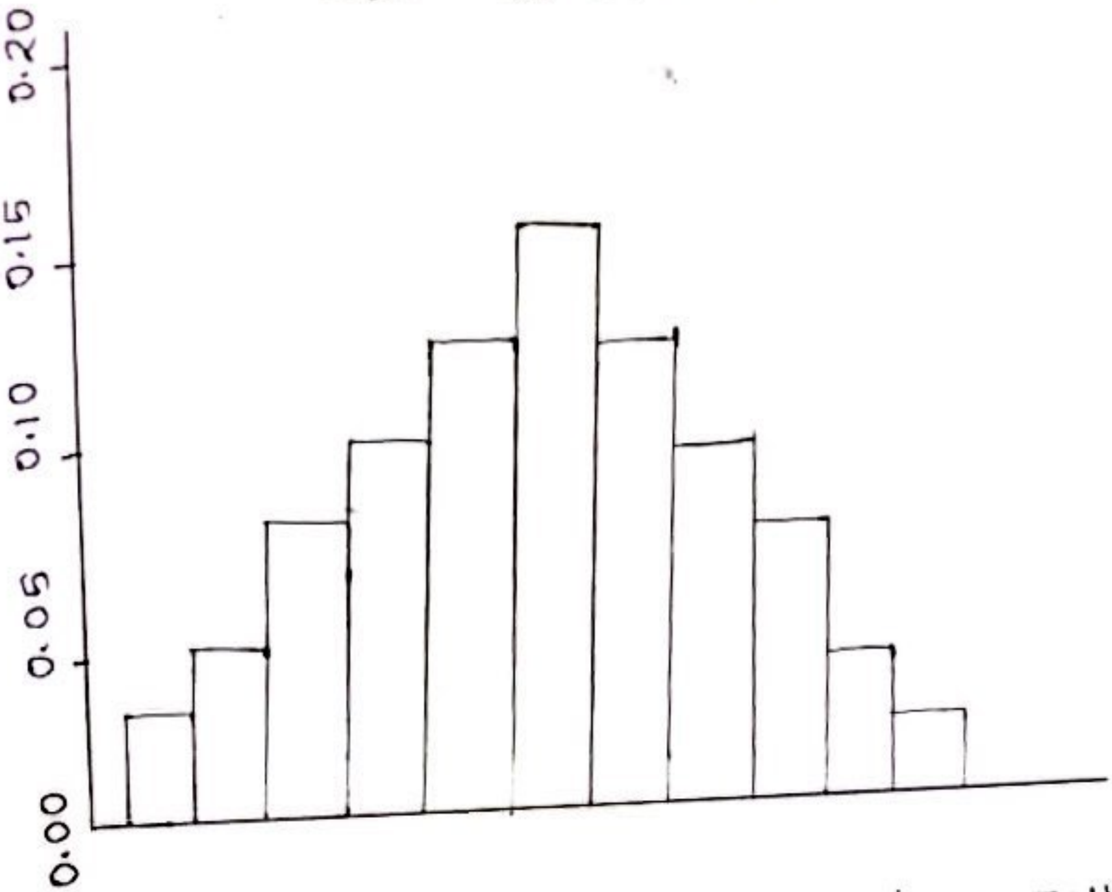
Input and output:

Es = 7

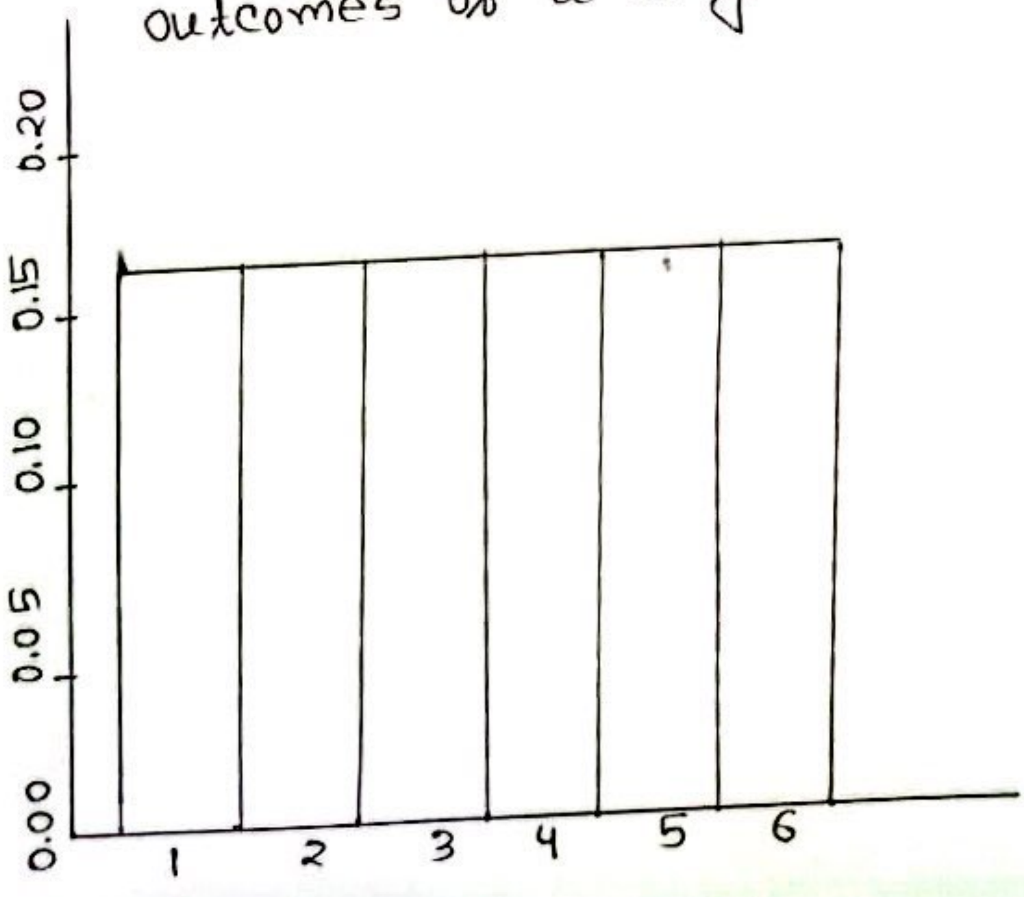
vars = 5.833

Scanned with CamScanner

Sum of two dice rolls



outcomes of a single dice roll



Scanned with CamScanner

Comment : Two dic rolled, S is the sum of both faces, the expectation of S, $E(S) = 7$ and variance of S, $V(S) = 5.833$

Scanned with CamScanner

Experiment no: 09

Name of the experiment: The number of systolic blood pressure of healthy subjects. The data set contains $n = 25$.

120, 115, 94, 118, 111, 102, 102, 131, 104, 107, 115, 139, 115, 113, 114, 103, 115, 134, 109, 109, 93, 118, 109, 106, 125

Do you think that the sample follows $N(\mu, 100)$.

Objectives:

1. To calculate the variance test
2. To calculate null hypothesis
3. To comment on data.

Procedure:

1. First we have to select the null hypothesis and alternative hypothesis.

2. Then we have to select the level of significance.

3. Then we need to select the test statistics.

4. Then we have to formulate the decision rule.

Scanned with CamScanner

Experiment No : 12

Experiment Name : Test the hypothesis that the median systolic blood pressure of healthy subjects (status-0) and Subject with hypertension (status-1) are equal have $d_0 = 0$. The dataset contains $n_1 = 25$ subjects with status-0 and $n_2 = 30$ with status-1.

status-0 : (120, 115, 94, 118, 111, 102, 102, 131, 104, 107, 115, 139, 115, 113, 114, 105, 115, 134, 109, 109, 93, 118, 109, 106, 125)

Status-1 : (150, 142, 119, 127, 141, 149, 144, 142, 149, 161, 143, 140, 148, 149, 141, 146, 159, 152, 135, 134, 161, 130, 125, 141, 148, 153, 145, 137, 147, 169)

Is there any difference in the median between status-0 and status-1?

Scanned with CamScanner

Objectives:

1. To calculate the difference in the median between status-0 and status-1.
2. To calculate p-value.
3. To comment on the data.

Procedure:

Step-1: Select the null hypothesis and alternate hypothesis. The null hypothesis state that there is no difference in the median between status-0 and status-1. The alternate hypothesis state that there is difference in the median between status-0 and status-1.

$$H0: md1 = md2$$

$$H1: md1 \neq md2$$

Scanned with CamScanner

Step-2: Select the level of significance.
The selected level of significance is 0.05.

Step-3: Select the test statistics.
There are two valued non parametric so the test statistics is wilcoxon rank sum test.

Step-4: Formulate the decision rule.
If P value is greater than alpha then the null hypothesis is accepted otherwise null hypothesis is rejected.

R-Source Code:

```
X1<-c(120,115,94,118,111,102,102,131,104,107,
115,139,115,114,113,105,115,134,109,109,93,
118,109,106,125)
X2<-c(150,142,119,127,141,149,144,142,149,161,
143,140,148,149,141,146,159,152,135,134,161,130,
125,141,148,153,145,137,147,169)
```

Scanned with CamScanner

```
wilcox.test(X1,X2, exact=FALSE,
correct=TRUE, alternative="two.sided")
```

Input and output:

W=18.

p-value = 1.649×10^{-9}

Comment: From the R code we can see that, p-value is less than alpha. p-value < alpha, so the null hypothesis is rejected. We can say that, There is difference in the median between status-0 and status-1.

Scanned with CamScanner