Case Study – Statistical Inference

4. Statistical Inference

4.1 Parametric Testing of Hypothesis

4.1.1 Testing for Mean of a Normal Population

4.1.1.1 When Variance is known

Case 1: Based on the following sample of size 10 from a Normal Population (with variance = 3) of Monthly Exenditure Students of Delhi University: 5000, 2000, 3000, 3456, 3623, 5200, 3400, 1200, 4500, 3500, I want to test the following hypotheses:

i. Average Monthly Expenditure of the students of Delhi University is 3500 or not.
ii. Average Monthly Expenditure of the students of Delhi University is 3500 or more than that.

Methodology:

We do z test for the above problem.

R Code and Output:

```
#case 1(var is known)

DU<-c(5000, 2000, 3000, 3456, 3623, 5200, 3400, 1200, 4500, 3500)

smean<-mean(DU)


install.packages("BSDA")

library(BSDA)

z.test(DU,mu=3500,sigma.x=sqrt(3),alternative="two.sided")

z.test(DU,mu=3500,sigma.x=sqrt(3),alternative="greater")

pnorm(smean,mean=3500,sd=sqrt(3),lower.tail = T)<(0.05/2) #two tailed

pnorm(smean,mean=3500,sd=sqrt(3),lower.tail = F)<0.05
```

Output:

```
DU<-c(5000, 2000, 3000, 3456, 3623, 5200, 3400, 1200, 4500, 3500)
> smean<-mean(DU)
> z.test(DU,mu=3500,sigma.x=sqrt(3),alternative="two.sided")

        One-sample z-Test

data:  DU
z = -22.091, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 3500
95 percent confidence interval:
 3486.826 3488.974
sample estimates:
mean of x
   3487.9

> z.test(DU,mu=3500,sigma.x=sqrt(3),alternative="greater")

        One-sample z-Test

data:  DU
z = -22.091, p-value = 1
alternative hypothesis: true mean is greater than 3500
95 percent confidence interval:
 3486.999       NA
sample estimates:
mean of x
   3487.9

>
> pnorm(smean,mean=3500,sd=sqrt(3),lower.tail = F)<0.05
[1] FALSE
> pnorm(smean,mean=3500,sd=sqrt(3),lower.tail = T)<(0.05/2)
[1] TRUE
```

Conclusion:

i) Average monthly expenditure of DU students is not 3500
ii) Average monthly expenditure of DU students is not greater than 3500

4.1.1.2. When Variance is unknown

Case 2: Based on the following sample of size 10 from a Normal Population (variance is unknown) of Monthly Expenditure Students of Delhi University: 5000, 2000, 3000, 3456, 3623, 5200, 3400, 1200, 4500, 3500, I want to test the following hypotheses:

- **i.** Average Monthly Expenditure of the students of Delhi University is 3500 or not.
- **ii.** Average Monthly Expenditure of the students of Delhi University is 3500 or more than that.

R Code and Output:

```
#var is unknown

pt(smean,df=9,lower.tail=T)<(0.05/2)

t.test(DU,mu=3500,alternative="two.sided")


pt(smean,df=9,lower.tail=F)<0.05

t.test(DU,mu=3500,alternative="greater")
```

Output:

```
> pt(smean,df=9,lower.tail=T)<(0.05/2)
[1] FALSE
> t.test(DU,mu=3500,alternative="two.sided")

        One Sample t-test

data:  DU
t = -0.030728, df = 9, p-value = 0.9762
alternative hypothesis: true mean is not equal to 3500
95 percent confidence interval:
 2597.109 4378.691
sample estimates:
mean of x
   3487.9
> pt(smean,df=9,lower.tail=F)<0.05
[1] TRUE
> t.test(DU,mu=3500,alternative="greater")

        One Sample t-test

data:  DU
t = -0.030728, df = 9, p-value = 0.5119
alternative hypothesis: true mean is greater than 3500
95 percent confidence interval:
 2766.058      Inf
sample estimates:
mean of x
   3487.9
```

Conclusion:

i) Yes, average monthly expenditure of DU students is 3500
ii) No, , average monthly expenditure of DU students is not greater than 3500

Case 3: The mean weekly sales of soap bars in a chain of departmental stores was 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the campaign successful?

Methodology:

R code:

```
#case3

teststat<-(sqrt(22)*(153.7-146.3))/17.2

pnorm(teststat,lower.tail = F)
```

Output:

```
pnorm(teststat,lower.tail = F)
[1] 0.02179722
```

Conclusion:

No, the campaign was not successful

4.1.2. TESTING FOR VARIANCE OF A NORMAL POPULATION

Case 4: It is believed that the precision (as measured by the variance) of an instrument is no more than 0.16. Write down the null and alternative hypothesis for testing this belief. Carry out the test at 1% level given 11 measurements of the sample subject on the instrument: 2.5, 2.3, 2.4, 2.3, 2.5, 2.6, 2.5, 2.6, 2.6, 2.7, 2.5.

Methodology:

R code:

```
#case4

mea<-c(2.5, 2.3, 2.4, 2.3, 2.5, 2.6, 2.5, 2.6, 2.6,2.7, 2.5)

variance<-1/0.16

teststatistics=(11-1)*var(mea)/variance

pchisq(teststatistics,df=10,lower.tail=T)<0.01
```

Output:

```
> #case4
> mea<-c(2.5, 2.3, 2.4, 2.3, 2.5, 2.6, 2.5, 2.6, 2.6,2.7, 2.5)
> variance<-1/0.16
> teststatistics=(11-1)*var(mea)/variance
> pchisq(teststatistics,df=10,lower.tail=T)<0.01
[1] TRUE
```

Conclusion:

Yes, the precision is indeed no more than 0.16

4.1.3. Testing for Equality of Means of two Normal Populations

4.1.3.1. When Population Variances are known

Case 5: The means of two samples of 1000 and 2000 members (from Normal Population) are 67.5 inches and 68.9 inches respectively. Can the samples be regarded as drawn from the same population of standard deviation 2.5 inches?

Methodology:

R code:

```
#case5

tstat<-(67.9-68.9)/(sqrt((1/1000)+(1/2000))*2.5)

pnorm(tstat,lower.tail = T)<(0.05/2)
```

output:

```
#case5
>
> tstat<-(67.9-68.9)/(sqrt((1/1000)+(1/2000))*2.5)
> pnorm(tstat,lower.tail = T)<(0.05/2)
[1] TRUE
```

Conclusion:

No, they cant be regarded to be coming from same population at 0.05 level of significance

4.1.3.2. When Population Variances are unknown

Case 6: The heights of six randomly chosen sailors are (in inches): 63, 65, 68, 69, 71, 72 and those of 10 randomly chosen soldiers are: 61, 62, 65, 66, 69, 69, 70, 71, 72, and 73. Assuming that the samples are coming from Normal populations test if sailors are on the average taller than soldiers.

Methodology:

R code:
#var not given
sailors<-c(63, 65, 68, 69, 71, 72)
soldiers<-c(61, 62, 65, 66, 69, 69, 70, 71, 72,73)
t.test(sailors,soldiers,alternative = "two.sided")

output:

```
#var not given
> sailors<-c(63, 65, 68, 69, 71, 72 )
> soldiers<-c(61, 62, 65, 66, 69, 69, 70, 71, 72,73)
> t.test(sailors,soldiers,alternative = "two.sided")

        Welch Two Sample t-test

data:  sailors and soldiers
t = 0.10388, df = 12.228, p-value = 0.9189
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.98616  4.38616
sample estimates:
mean of x mean of y 
     68.0      67.8 
```

Conclusion:

No, on average the sailors are of same height as soldiers.

4.1.3.3. When Two Samples are Related

Case 7: Consider the *sleep* data available in R. Under the assumption of normality, can we say that the effect of the two soporific drugs is same?

Methodology:

R code:

```
#related data(case 7)
v<-sleep
t.test(v$extra[1:10],v$extra[11:20],alternative = "two.sided",paired = T)
```

Output:

```
#related data(case 7)
> v<-sleep
> t.test(v$extra[1:10],v$extra[11:20],alternative = "two.sided",paired = T)

        Paired t-test

data:  v$extra[1:10] and v$extra[11:20]
t = -4.0621, df = 9, p-value = 0.002833
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.4598858 -0.7001142
sample estimates:
mean of the differences
                  -1.58
```

Conclusion:

No, the effect of two soporific drugs is not same.

4.1.4. Testing for Equality of Variances of two Normal Populations

Case 8: Verify equality of variance in Case 6

Methodology:

R code:

```
#case8
var.test(v$extra[1:10],v$extra[11:20])
```

Output:

```
var.test(v$extra[1:10],v$extra[11:20])

        F test to compare two variances

data:  v$extra[1:10] and v$extra[11:20]
F = 0.79834, num df = 9, denom df = 9, p-value = 0.7427
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.198297 3.214123
sample estimates:
ratio of variances
         0.7983426
```

Conclusion:

Yes, the variance of two variables on specific population is same.

4.1.5. TESTING FOR EQUALITY OF MEANS OF SEVERAL NORMAL POPULATIONS

Case 9: Consider the *ChickWeight* data available in R. Under the assumption of normality, can we say that all the different diets have same effect on the weight of Chicken? If not, then find the best one for mass production.

Methodology:

R code:

```
#case9

View(ChickWeight)
fit <- aov(weight ~ Diet, data=ChickWeight)
summary(fit)
tapply(ChickWeight$weight,ChickWeight$Diet,"mean")
```

output:

```
fit <- aov(weight ~ Diet, data=ChickWeight)
> summary(fit)
             Df  Sum Sq Mean Sq F value   Pr(>F)    
Diet          3  155863   51954   10.81 6.43e-07 ***
Residuals   574 2758693    4806                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> tapply(ChickWeight$weight,ChickWeight$Diet,"mean")
       1        2        3        4 
102.6455 122.6167 142.9500 135.2627 
```

Conclusion:

No, different diets do not have same effect on chicken’s weight.
Diet 3 is the best for production.

4.2. NON-PARAMETRIC TESTING OF HYPOTHESIS

4.2.1. TESTING FOR POPULATION MEDIAN – FRANK WILCOXON ONE-SAMPLE SIGN TEST

Case 10: The following table represents observations on heights and weights of 15 females:

Obs	Height (in inches)	Weight (in lbs)	Obs	Height (in inches)	Weight (in lbs)
1	58	115	9	66	139
2	59	117	10	67	142
3	60	120	11	68	146
4	61	123	12	69	150
5	62	126	13	70	154
6	63	129	14	71	159
7	64	132	15	72	164
8	65	135			

Use sign test to test the following two hypothesis:

1. The Height of the females can be taken to be equal to 64 inches.
2. The Weight of the females can be taken to be equal to 135 lbs.

Methodology:
R code:

```
#case10
height<-c(58,59,60,61,62,63,64,65,66,67,68,69,70,71,72)
weight<-c(115,117,120,123,126,129,132,135,139,142,146,150,154,159,164)
med<-median(height)
#height
sum<-sum(height>med)
pbinom(sum,size=length(height),p=0.5,lower.tail = F)<(0.05/2)
#weight
med1<-median(weight)
sum1<-sum(weight>med1)
pbinom(sum1,size=length(weight),p=0.5,lower.tail = F)<(0.05/2)
```

Output:

```
> pbinom(sum,size=length(height),p=0.5,lower.tail = F)<(0.05/2)
[1] FALSE
> pbinom(sum1,size=length(weight),p=0.5,lower.tail = F)<(0.05/2)
[1] FALSE
```

Conclusion:

Yes, the Height of the females can be taken to be equal to 64 inches.

Yes, weight of the females can be taken to be equal to 135 lbs

Case 11: Win/Loss records of a certain basketball team during their 50 consecutive games are given in the following table:

Game	Outcome	Game	Outcome	Game	Outcome	Game	Outcome	Game	Outcome
1	1	11	1	21	0	31	0	41	1
2	1	12	1	22	1	32	1	42	0
3	1	13	1	23	1	33	1	43	0
4	1	14	0	24	1	34	1	44	0
5	1	15	1	25	1	35	1	45	1
6	1	16	0	26	0	36	1	46	1
7	0	17	1	27	1	37	1	47	0
8	1	18	1	28	1	38	0	48	1
9	1	19	1	29	1	39	0	49	1
10	1	20	0	30	0	40	1	50	1

Using Sign Test to test the hypothesis that win and loss are equally likely.

Methodology:

R code:

```
games<-c(1:50)
outcomes<-c(1,1,1,1,1,1,0,1,1,1,1,1,1,0,1,0,1,1,1,0,0,1,1,
            1,1,0,1,1,1,0,0,1,1,1,1,1,1,0,0,1,1,0,0,0,1,1,0,1,1,1)
data<-cbind(games,outcomes)
positives<-sum(outcomes)
negatives<-50-positives
library(BSDA)
SIGN.test(positives,negatives,alternative = "two.sided")
```

Output:

```
SIGN.test(positives,negatives,alternative = "two.sided")

        Dependent-samples Sign-Test

data:  positives and negatives
S = 1, p-value = 1
alternative hypothesis: true median difference is not equal to 0
0 percent confidence interval:
 22 22
sample estimates:
median of x-y
           22
```

Conclusion:

Yes, win and loss are equally likely indeed!

4.2.2. TESTING FOR EQUALITY OF TWO POPULATIONS

4.2.2.1. WALD-WOLFOWITZ RUN TEST

4.2.2.2. MANN-WHITNEY-WILCOXON U TEST

Case 12: An experiment on reading ability of students was conducted, where at the beginning of the year a class was randomly divided into two groups. One group was taught to read using a uniform method, where all the students progressed from one stage to the next at the same time, following the instructor's direction. The second group was taught to read using an individual method, where each student progressed at his own rate according to a programmed work book under the supervision of the instructor. At the end of the year each student was given a reading ability test and following were their scores:

First Group	227	176	252	149	16	55	234	194	247	92	184	147	88	161	171
Second Group	202	14	165	171	292	271	151	235	147	99	63	284	53	228	271

Use U-test to test if two different teaching methods for reading ability can be taken as equally effective.

Methodology:

R code:

```
#case12
first_group<-c(227,176,252,149,16,55,234,194,247,92,184,147,88,161,171)
second_group<-c(202,14,165,171,292,271,151,235,147,99,63,284,53,228,271)

wilcox.test(first_group,second_group,alternative = "two.sided")
```

output:

```
wilcox.test(first_group,second_group,alternative = "two.sided")

        Wilcoxon rank sum test with continuity correction

data:  first_group and second_group
W = 97, p-value = 0.5337
alternative hypothesis: true location shift is not equal to 0
```

conclusion:

Yes, two different teaching methods for reading ability can be taken as equally effective

4.2.2.3. KOLMOGOROV-SMIRNOV TEST

Case 13: For the following two samples test if they can be taken to be coming from same population:

Observation	Sample 1	Sample 2
1	0.075204597	1.319177696
2	0.282203071	0.255423126
3	0.473605304	0.250284353
4	0.171775727	0.941835437
5	0.084642496	3.078396099
6	0.601160542	0.270368067
7	0.212552515	0.413272132
8	0.294969478	0.05425652
9	0.026919861	1.340734424
10	0.054462148	0.127618122
11	0.076084169	0.060699583
12	0.021943532	0.208278913
13	0.486042232	0.104869289
14	0.083376869	1.126610877
15	0.62800881	1.179774988
16	1.317637268	2.015836491
17	0.431532897	0.43267859
18	0.151809043	0.686019322
19	0.645182388	1.210587738
20	0.018898663	0.230682213

Methodology:
R code:

```
#case13
sample1<-c(0.075204597,0.282203071,0.473605304,0.171775727,0.084642496,
           0.601160542,0.212552515,0.294969478,0.026919861,0.054462148,
           0.076084169,0.021943532,0.486042232,0.083376869,0.62800881,
           1.317637268,0.431532897,0.151809043,0.645182388,0.018898663)
sample2<-c(1.319177696,0.255423126,0.250284353,0.941835437,3.078396099,
           0.270368067,0.413272132,0.05425652,1.340734424,0.127618122,
           0.060699583,0.208278913,0.104869289,1.126610877,1.179774988,
           2.015836491,0.43267859,0.686019322,1.210587738,0.230682213)
ks.test(sample1,sample2,alternative = "two.sided")
```

Output:

```
ks.test(sample1,sample2,alternative = "two.sided")

        Two-sample Kolmogorov-Smirnov test

data:  sample1 and sample2
D = 0.4, p-value = 0.08106
alternative hypothesis: two-sided
```

Conclusion: Yes, they can be considered to be coming from same population.

4.2.2.4. Two-Sample Sign Test

Case 14: Following data represents the marks given to the same set 22 students by two different professors in the same examination:

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Professor A	79	87	24	41	59	12	91	78	63	30	09	64	50	92	64	39	49	86	23	45	12	88
Professor B	83	91	18	39	67	34	78	89	38	45	10	45	56	89	67	35	40	82	32	38	23	92

Using Sign Test, test if the grading of both the professors can be taken to be same.

Methodology:

R code:

```
#case14

prof1<-c(79,87,24,41,59,12,91,78,63,30,09,64,50,92,64,39,49,86,23,45,12,88)
prof2<-c(83,91,18,39,67,34,78,89,38,45,10,45,56,89,67,35,40,82,32,38,23,92)
SIGN.test(prof1,prof2,alternative = "two.sided")
```

Output:

```
SIGN.test(prof1,prof2,alternative = "two.sided")

        Dependent-samples Sign-Test

data:  prof1 and prof2
S = 10, p-value = 0.8318
alternative hypothesis: true median difference is not equal to 0
95 percent confidence interval:
 -6.139343  4.139343
sample estimates:
median of x-y
           -2

                  Conf.Level  L.E.pt U.E.pt
Lower Achieved CI     0.9475 -6.0000 4.0000
Interpolated CI       0.9500 -6.1393 4.1393
Upper Achieved CI     0.9831 -8.0000 6.0000
```

Conclusion:

Yes, the grading of two professors can be considered same.

4.2.3. Testing for Equality of several Populations – Kruskal Wallis Test

Case 15: Consider the *ChickWeight* data available in R. Under no assumption of normality, can we say that all the different diets have same effect on the weight of Chicken? If not, then find the best one for mass production.

Methodology:

R code:

```
#case15
View(ChickWeight)
kruskal.test(weight~Diet,data=ChickWeight)
```

output:

```
kruskal.test(weight~Diet,data=ChickWeight)

        Kruskal-Wallis rank sum test

data:  weight by Diet
Kruskal-Wallis chi-squared = 24.45, df = 3, p-value = 2.012e-05
```

Conclusion:

No, different diets do not have same effect on the weight of Chicken.

4.2.4. Goodness of Fit – Kolmogorov-Smirnov Test

Case 16: For the following four samples test if they are drawn from Normal, Exponential, Poisson and Uniform distributions respectively.

Observation	Sample 1	Sample 2	Sample 3	Sample 4
1	1.089781309	0.046136443	4	10.25068427
2	1.962787672	0.296905535	2	10.12379195
3	1.724451834	0.013852846	1	17.81733259
4	1.63955842	0.149763684	1	18.87337658
5	0.144050286	0.216846562	2	15.32347378
6	0.232942589	0.549152735	3	11.97729205
7	1.68271611	0.075868307	4	16.79090476
8	3.633887711	0.147932045	3	14.40535435
9	1.81341443	0.29035859	3	14.10547096
10	1.683039558	0.027180583	4	14.99055234
11	1.659612162	0.163903305	3	12.68408943
12	0.8396626	0.8104371	1	10.62609998
13	3.427254188	0.078686029	8	15.59840961
14	1.127955432	0.153359897	4	17.59452935
15	1.552543896	0.141724322	5	10.60249139
16	0.214796062	0.066255849	2	17.17324608
17	0.475882672	0.085298693	4	11.59441059
18	3.013061127	0.507875983	5	14.11860911
19	2.73502768	0.104899753	0	19.68738385
20	2.583921184	0.020127363	5	17.20303417

Methodology:

R code:

```
#case16
sample1<-c(1.089781309,1.962787672,1.724451834,1.63955842,0.144050286,0.232942589,
	1.68271611,3.633887711,1.81341443,1.683039558,1.659612162,0.8396626,
	3.427254188,1.127955432,1.552543896,0.214796062,0.475882672,3.013061127,
	2.73502768,2.583921184)
sample2<-c(0.046136443,0.296905535,0.013852846,0.149763684,0.216846562,0.549152735,
	0.075868307,0.147932045,0.29035859,0.027180583,0.163903305,0.8104371,
	0.078686029,0.153359897,0.141724322,0.066255849,0.085298693,0.507875983,
	0.104899753,0.020127363)
sample3<-c(4,2,1,1,2,3,4,3,3,4,3,1,8,4,5,2,4,5,0,5)
sample4<-c(10.25068427,10.12379195,17.81733259,18.87337658,15.32347378,11.97729205,
	16.79090476,14.40535435,14.10547096,14.99055234,12.68408943,10.62609998,
```

```
    15.59840961,17.59452935,10.60249139,17.17324608,11.59441059,14.11860911,
    19.68738385,17.20303417)
set.seed(100)
ks.test(sample1,rnorm(length(sample1)))
ks.test(sample2,rexp(length(sample2)))
ks.test(sample3,rpois(length(sample3),lambda=mean(sample3)))
ks.test(sample4,runif(length(sample4)))
```

Output:

```
ks.test(sample1,rnorm(length(sample1)))

        Two-sample Kolmogorov-Smirnov test

data:  sample1 and rnorm(length(sample1))
D = 0.7, p-value = 5.569e-05
alternative hypothesis: two-sided

> ks.test(sample2,rexp(length(sample2)))

        Two-sample Kolmogorov-Smirnov test

data:  sample2 and rexp(length(sample2))
D = 0.75, p-value = 9.547e-06
alternative hypothesis: two-sided

> ks.test(sample3,rpois(length(sample3),lambda=mean(sample3)))

        Two-sample Kolmogorov-Smirnov test

data:  sample3 and rpois(length(sample3), lambda = mean(sample3))
D = 0.15, p-value = 0.978
alternative hypothesis: two-sided

> ks.test(sample4,runif(length(sample4)))

        Two-sample Kolmogorov-Smirnov test

data:  sample4 and runif(length(sample4))
D = 1, p-value = 1.451e-11
alternative hypothesis: two-sided
```

Conclusion:

Only sample 3 comes from poisson population. And the rest do not come from their respective populations.

4.2.5. Independence of Attributes – Chi-Square Test

Case 17: Out of 8000 graduates in a town 800 are females, out of 1600 graduate employees 120 are females. Use Chi – Square test to test if any sex discrimination is made in the employment.

Methodology:

R code:

```
#case17
employfema<-matrix(c(800,7200,120,1480),nrow=2)
chisq.test(employfema)
```

Output:

```
chisq.test(employfema)

        Pearson's Chi-squared test with Yates' continuity correction

data:  employfema
X-squared = 9.331, df = 1, p-value = 0.002253
```

Conclusion:

Yes, sex discrimination is made in the employment.