LAB-11

Paired t-test And F- (Variance Ratio Test)

AIM: to analyse the improvement or effectiveness of a new methodology adopted. And also test the hypothesis for variance ratio.

HYPOTHESIS TESTS FOR MEAN DIFFERENCES: PAIRED DATA-t-TEST

Problem 1 :

A school athletics has taken a new instructor, and want to test the effectiveness of the new type of training proposed by the new instructor comparing the average times of 10 runners in the 100 meters. The results are given below(time in seconds)

Before training	*12.9*	*13.5*	*12.8*	*15.6*	*17.2*	*19.2*	*12.6*	*15.3*	*14.4*	*11.3*
After training	*12.7*	*13.6*	*12.0*	*15.2*	*16.8*	*20.0*	*12.0*	*15.9*	*16.0*	*11.1*

Solution: *In this case we have two sets of paired samples, since the measurements were made on the same athletes before and after the workout. To see if there was an improvement, deterioration, or if the means of times have remained substantially the same (hypothesis H0), we need to make a Student's t-test for paired samples, proceeding in this way*

```
> before = c(12.9, 13.5, 12.8, 15.6, 17.2, 19.2, 12.6, 15.3, 14.4, 11.3)
> after = c(12.7, 13.6, 12.0, 15.2, 16.8, 20.0, 12.0, 15.9, 16.0, 11.1)
> t.test(before,after, paired=TRUE)

        Paired t-test

data:  before and after
t = -0.21331, df = 9, p-value = 0.8358
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5802549  0.4802549
sample estimates:
mean of the differences
                  -0.05
```

Interpretation :-

The p-value is greater than 0.05, then we do not reject the hypothesis H_0 of equality of the averages and conclude that the new training has not made any significant improvement to the team of athletes.

Problem 2 :-

Suppose now that the manager of the team (given the results obtained) fired the coach who has not made any improvement, and take another, more promising. We report the times of athletes after the second training:

Before training:	*12.9*	*13.5*	*12.8*	*15.6*	*17.2*	*19.2*	*12.6*	*15.3*	*14.4*	*11.3*
After the second training:	*12.0*	*12.2*	*11.2*	*13.0*	*15.0*	*15.8*	*12.2*	*13.4*	*12.9*	*11.0*

R code:-

Now we check if there was actually an improvement, ie perform a t-test for paired data, specifying in R to test the alternative hypothesis H1 of improvement in times. To do this simply add the syntax alt = "less" when you call the t-test:

```
> before=c(12.9, 13.5, 12.8, 15.6, 17.2, 19.2, 12.6, 15.3, 14.4, 11.3)
> after = c(12.0, 12.2, 11.2, 13.0, 15.0, 15.8, 12.2, 13.4, 12.9, 11.0)
> t.test(before,after, paired=TRUE, alt="less")

        Paired t-test

data:  before and after
t = 5.2671, df = 9, p-value = 0.9997
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 2.170325
sample estimates:
mean of the differences
                   1.61
```

Interpretation:-

In response, we obtained a p-value well above 0.05, which leads us to conclude that we can reject the null hypothesis H_0 in favour of the alternative hypothesis H1: the new training has made substantial improvements to the team.

Problem 3 *: Consider the paired data below that represents cholesterol levels on 10 men before and after a certain medication*

Before(x)	*237*	*289*	*257*	*228*	*303*	*275*	*262*	*304*	*244*	*233*
After(y)	*194*	*240*	*230*	*186*	*265*	*222*	*242*	*281*	*240*	*212*

Test the claim that, on average, the drug lowers cholesterol in all men. I.e., test the claim that $\mu_d > 0$*. Test this at the 0.05 significance level.*

R-code:-

```
> before=c(237,289,257,228,303,275,262,304,244,233)
> after=c(194,240,230,186,265,222,242,281,240,212)
> t.test(before,after,paired=TRUE,alternative="greater",mu=0)

        Paired t-test

data:  before and after
t = 6.5594, df = 9, p-value = 5.202e-05
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 23.05711      Inf
sample estimates:
mean of the differences
                     32
```

Interpretation :-

We can reject the null hypothesis and support the claim because the P-value ($\approx 5.2X10^{-5}$) is less than the signicance level.

F Test to Compare Two Variances

Problem 1 :-

Five Measurements of the output of two units have given the following results (in kilograms of material per one hour of operation) .Assume that both samples have been obtained from normal populations, test at 10% significance level if two populations have the same variance.

Unit A	*14.1*	*10.1*	*14.7*	*13.7*	*14.0*
Unit B	*14.0*	*14.5*	*13.7*	*12.7*	*14.1*

R code:

$H_0: S_1^2 = S_2^2$

$H_1: S_1^2 \neq S_2^2$

Level of Significance :0.10

R code:-

```
> Unit_A=c(14.1,10.1,14.7,13.7,14.0)
> Unit_B=c(14.0,14.5,13.7,12.7,14.1)
> var.test(Unit_A,Unit_B)

        F test to compare two variances

data:  Unit_A and Unit_B
F = 7.3304, num df = 4, denom df = 4, p-value = 0.07954
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
  0.7632268 70.4053799
sample estimates:
ratio of variances
          7.330435
```

Inference : here p valuve >0.05 ,then there is no evidence to reject the null hypothesis.

Problem 2: Energy Data :- (Variance Ratio-test)

```
> energy=read.csv("C:\\Users\\aadmin\\Desktop\\energy.csv")
> energy
   expend stature
1    9.21   obese
2    7.53    lean
3    7.48    lean
4    8.08    lean
5    8.09    lean
6   10.15    lean
7    8.40    lean
8    0.88    lean
9    6.13    lean
10   7.90    lean
11  11.51   obese
12  12.79   obese
13   7.05    lean
14  11.85   obese
15   9.97   obese
16   7.48    lean
17   8.79   obese
18   9.69   obese
19   9.68   obese
20   7.58    lean
21   9.19   obese
22   8.11    lean
> var.test(energy$expend~energy$stature)

        F test to compare two variances

data:  energy$expend by energy$stature
F = 2.321, num df = 12, denom df = 8, p-value = 0.2386
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5526712 8.1509583
sample estimates:
ratio of variances
          2.321035
```

Inference :

Here p valuve >0.05 ,then there is no evidence to reject the null hypothesis.

Practice questions:-

1. *A study was performed to test whether cars get better mileage on premium gas than on regular gas. Each of 10 cars was first filled with either regular or premium gas, decided by a coin toss, and mileage for that tank was recorded. The mileage was recorded again for the same car using the other kind of gasoline. We use a paired t – test to determine whether cars get significant better mileage with premium gas.*

Regular	***16***	***20***	***21***	***22***	***23***	***22***	***27***	***25***	***27***	***28***
Premium	*19*	*22*	*24*	*24*	*25*	*25*	*26*	*26*	*28*	

2. *The Scores of 10 candidates prior and after training are given below*

Prior	*84*	*48*	*36*	*37*	*54*	*69*	*83*	*96*	*90*	*65*
After	*90*	*58*	*56*	*49*	*62*	*81*	*84*	*86*	*84*	*75*

Test the training is Effective or Not?

3. *An IQ test was administrated to 5 persons before and after they were trained. The results are given below*

Candidates	*I*	*II*	*III*	*IV*	*V*
IQ before Training	*110*	*120*	*123*	*132*	*125*
IQ After Training	*120*	*118*	*125*	*136*	*121*

Test whether there is any change in IQ after the training Programme

4. *In order to compare the effectiveness of two sources of nitrogen, namely ammonium chloride and urea on grain yield of paddy, an experiment was conducted. The results on the grain yield of paddy(kg/plot) under the two treatments are given below*

Ammonium chloride	*13.4*	*10.9*	*11.2*	*11.8*	*14.0*	*15.3*	*14.2*	*12.6*	*17.0*	*16.2*	*16.5*	*15.7*
Urea	*12.0*	*11.7*	*10.7*	*11.2*	*14.8*	*14.4*	*13.9*	*13.7*	*16.9*	*16.0*	*15.6*	*16.0*

Asses which sources nitrogen is better for paddy

5. In order to compare the effectiveness of two sources of nitrogen, namely ammonium chloride and urea on grain yield of paddy, an experiment was conducted. The results on the grain yield of paddy(kg/plot) under the two treatments are given below

Ammonium chloride	13.4	10.9	11.2	11.8	14.0	15.3	14.2	12.6	17.0	16.2	16.5	15.7
Urea	12.0	11.7	10.7	11.2	14.8	14.4	13.9	13.7	16.9	16.0	15.6	16.0

Asses which sources nitrogen is better for paddy