Statistics for Engineers (MAT2001)-Lab Experiment-VIII: Small sample test

1 t-test

The t.test( ) function produces a variety of t-tests. Unlike most statistical packages, the default assumes unequal variance.

- one sample t-test: t.test(y,mu=mean value)
- independent 2-group t-test: t.test(y x) where y is numeric and x is a binary factor
- independent 2-group t-test: t.test(y1,y2) where y1 and y2 are numeric
- paired t-test: t.test(y1,y2,paired=TRUE) where y1 and y2 are numeric

An outbreak of salmonella-related illness was attributed to ice produced at a certain factory. Scientists measured the level of Salmonella in 9 randomly sampled batches ice crean.The levels(in MPN/g) were: 0.593, 0.142, 0.329, 0.691, 0.231, 0.793, 0.519, 0.392, 0.418. Is there evidence that the mean level pf Salmonella in ice cream greater than 0.3 MPN/g?

```
x <- c(0.593,0.142,0.329,0.691,0.231,0.793,0.519,0.392,0.418)
t.test(x, alternative = "greater", mu=0.3)

##
##  One Sample t-test
##
## data:  x
## t = 2.2051, df = 8, p-value = 0.02927
## alternative hypothesis: true mean is greater than 0.3
## 95 percent confidence interval:
##  0.3245133       Inf
## sample estimates:
## mean of x
## 0.4564444
```

From the output we see that the p-value = 0.029. Hence, there is moderately strong evidence that the mean Salmonella level in the ice cream is above 0.3 MPN/g.

Suppose that 10 volunteers have taken an intelligence test; here are the results obtained. The average score of the entire population is 75 in the same test. Is there any significant difference (with a significance level of 95%) between the sample and population means, assuming that the variance of the population is not known. Scores: 65, 78, 88, 55, 48, 95, 66, 57, 79, 81

```
scr <- c(65, 78, 88, 55, 48, 95, 66, 57, 79, 81)
t.test(scr, mu=75)

##
##  One Sample t-test
##
## data:  scr
## t = -0.78303, df = 9, p-value = 0.4537
## alternative hypothesis: true mean is not equal to 75
## 95 percent confidence interval:
##  60.22187 82.17813
## sample estimates:
## mean of x
##      71.2

qt(0.975, 9)

## [1] 2.262157
```

The t-computed value is smaller than t-tabulated, we accept the null hypothesis of equality of the averages.
Alternatively we could consider the p-value with a significance level of 95%. If p-value is greater than 0.05 then we accept the null hypothesis H0, otherwise we reject the null .

Comparing two independent sample means, taken from two populations with unknown variance.The following data shows the heights of individuals of two different countries with unknown population variances. Is there any significant difference b/n the average heights of two groups.
A: 175 168 168 190 156 181 182 175 174 179
B: 185 169 173 173 188 186 175 174 179 180

```
a <-c(175, 168, 168, 190, 156, 181, 182, 175, 174, 179)
b <- c(185, 169, 173, 173, 188, 186, 175, 174, 179, 180)
t.test(a, b, var.equal = FALSE, paired=FALSE)

##
##  Welch Two Sample t-test
##
## data:  a and b
```

```
## t = -0.94737, df = 15.981, p-value = 0.3576
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.008795   4.208795
## sample estimates:
## mean of x mean of y
##     174.8     178.2
```

```
qt(0.975, 10.224)
```

```
## [1] 2.221539
```

The p-value > 0.05, we conclude that the means of the two groups are significantly similar.
The value of t is less than the tabulated t-value for 10.224 df, we accept H0.
Suppose the recovery time for patients taking a new drug is measured (in days). A placebo group is also used to avoid the placebo effect. The data are as follows
with drug : 15 10 13 7 9 8 21 9 14 8
placebo : 15 14 12 8 14 7 16 10 15 2
Is there any significant difference between the average effect of these two drugs?

```
wd <- c(15, 10, 13, 7, 9, 8, 21, 9, 14, 8)
plc <- c(15, 14, 12, 8, 14, 7, 16, 10, 15, 2)
t.test(wd, plc,alternative="less",  var.equal=TRUE )
```

```
##
##  Two Sample t-test
##
## data:  wd and plc
## t = 0.050529, df = 18, p-value = 0.5199
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 3.531811
## sample estimates:
## mean of x mean of y
##      11.4      11.3
```

P value(0.5199) > 0.05 then there is no evidence to reject our Null hypothesis.

2 F-Test (Variance Ration Test)

`var.test(x, y)`

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, 14.5, 13.7, 12.7, 14.1

```
ua <- c(14.1, 10.1, 14.7, 13.7, 14.0)
ub <- c(14, 14.5, 13.7, 12.7, 14.1)
var.test(ua, ub)

##
##  F test to compare two variances
##
## data:  ua and ub
## 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
```

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

Practice Problem

1. A certain stimulus administered to each of the 13 patients resulted in the following increase of blood pressure: 5, 2, 8,-1, 3, 0, -2, 1, 5, 0, 4, 6, 8. Can it be concluded that the stimulus, in general, be accompanied by an increase in the blood pressure?

2. The manufacturer of a certain make of electric bulbs claims that his bulbs have a mean life of 25 months with a standard deviation of 5 months. Random samples of 6 such bulbs have the following values: Life of bulbs in months: 24, 20, 30, 20, 20, and 18. Can you regard the producer's claim to valid at 1% level of significance?

3. Data on weight (grams) of two treatments of NMU (nistroso- methyl urea) are recorded. Find out whether these two treatments have identical effects by using t test for sample means at 5% level of significance.
   Sample: 1 2 3 4 5 6 7 8 9 10 11 12
   Treatments 0.2 %: 2.0 2.7 2.9 1.9 2.1 2.6 2.7 2.9 3.0 2.6 2.6 2.7
   0.4% : 3.2 3.6 3.7 3.5 2.9 2.6 2.5 2.7

4. Hypothesis Tests for Two Means :Independent Data:Here we test for a difference in means for the following data
   No Drug(x1): 237 289 257 228 303 275 262 304 244 233,
   Drug(x2): 194 240 230 186 265 222 242 281 240 212