How to apply R in simple hypothesis testing

Hypothesis testing has been widely applied in many statistical analyses. Simple hypothesis testing compare differences between two samples by correlating them on their variances and mean values. We shall see how R works in these circumstances.

Using F-test

F-test studies any significant difference between the variances of two samples. For example:

```
> #F test
> Y=c(110,113,108,115,112)
> X=c(98,110,118,128,112,105,112)
> var.test(X,Y)

        F test to compare two variances

data:  X and Y
F = 12.348, num df = 6, denom df = 4, p-value = 0.02938
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
  1.342603 76.895082
sample estimates:
ratio of variances
          12.34834
```

As the p-value is smaller than 0.05, we conclude that we should reject Ho, indicating that there is a significant difference between the variances of X and Y.

Using the Student's t-test

The Student's t-test is a method to compare two samples, looking at the means to determine if the samples are different. This is a parametric test and the data should be normally distributed.

There are several versions of the t-test and R can handle these using the t.test( ) command which has a variety of options.

For example:

One sample t-testing against a certified or assigned value ‘mu’

```
> # single sample t test
> A=c(95,97,94,89,97,102,96)
> mean(A)
[1] 95.71429
> t.test(A,mu=100)

        One Sample t-test

data:  A
t = -2.9047, df = 6, p-value = 0.02717
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
 92.10406 99.32451
sample estimates:
mean of x
 95.71429
```

Since the p-value is smaller than 0.05, indicating that the Ho is rejected suggesting that the mean is significantly different from the given value mu=100.

For another set of data, the null hypothesis Ho is not rejected because the p-value is larger than 0.05.

```
> # single sample t test
> A=c(99,97,94,104,97,102,96)
> mean(A)
[1] 98.42857
> t.test(A,mu=100)

        One Sample t-test

data:  A
t = -1.1862, df = 6, p-value = 0.2804
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
  95.1869 101.6702
sample estimates:
mean of x
 98.42857
```

Two sample t-test with equal variance

We can override the default and use the classic t-test by adding the var.equal=TRUE instruction which forces the command to assume that the variance of the two samples is equal. The calculation of the t-value uses pooled variance and the degrees of freedom are unmodified.

For example:

```
> #2-sample t test
> P=c(95,97,89,97,102,96)
> Q=c(101,96,98,104,97,93)
> var.test(P,Q)

        F test to compare two variances

data:  P and Q
F = 1.1759, num df = 5, denom df = 5, p-value = 0.8632
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1645513 8.4037630
sample estimates:
ratio of variances
          1.175947

> t.test(P,Q,var.equal=TRUE)

        Two Sample t-test

data:  P and Q
t = -0.93, df = 10, p-value = 0.3743
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.357694  3.024360
sample estimates:
mean of x mean of y
 96.00000  98.16667
```

From the p-value obtained in the above two-sample t-test, we conclude that there is no significant difference between the two means, i.e. Ho is true.