STAT 40001/STAT 50001 Statistical Computing

Lecture 9

Department of Mathematics and Statistics

Hypothesis Testing

Most statistical inference centers around the parameters of a population(often the mean, standard deviation). Methods of drawing inferences about parameters are of two types:

- Actually estimate the value of the parameters
- Make decisions concerning the value of the parameters

When we estimate the value of the parameters we are using the method of estimations. Decisions concerning the value of the parameter are obtained by hypothesis testing.

The hypothesis testing procedure is a method for choosing between two competing hypotheses the so-called null hypothesis (H_o) and alternative hypothesis(H_a).

Null Vs. Alternative Hypotheses

Null Hypothesis (H_o): A null hypothesis is a claim (or statement) about a population parameter that is assumed to be true until it is declared false.

Alternative Hypothesis(H_a): An alternative hypothesis is a claim about population parameters that will be true if the null hypothesis is false.

Types of Statistical Tests

A statistical test is
i) left- tailed if H_a states that the parameter is less than the value claimed in H_o

$$H_0 : \mu \geq \mu_0$$
$$H_a : \mu < \mu_0$$

ii) right- tailed if H_a states that the parameter is greater than the value claimed in H_o

$$H_0 : \mu \leq \mu_0$$
$$H_a : \mu > \mu_0$$

iii) two- tailed if H_a states that the parameter is different from the value claimed in H_o.

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

Guidelines for Hypothesis Testing

- When testing a hypothesis concerning the value of some parameter, the statement of equality will always be included in H_o.
- Whatever is to be detected or supported is the alternative hypothesis.
- Since our research hypothesis is H_a, it is hoped that the evidence leads us to reject H_o and thereby to accept H_a

Hypothesis Testing

Once a sample has been selected and the data have been collected, a decision must be made. The decision will be one of the following

- Reject H_o
- Fail to reject H_o

Definition: The decision is made by observing the value of some statistic whose probability distribution is known under assumption that the null value is the true value of the parameter. Such statistic is called the test statistic.

Definition: A rejection region is the region that specifies the values of the observed test statistic for which the null hypothesis will be rejected.

Performing Hypothesis Tests

Few steps to perform hypothesis test

- Determine the parameter of interest.
- Determine the null hypothesis H_0.
- Determine the alternative hypothesis H_a
- Choose the appropriate test statistic.
- Determine the rejection region using the significance level.
- Determine if the test statistic falls into the rejection region or not.
- Draw your conclusion:
  - Reject the null hypothesis
  - Fail to reject the null hypothesis

Types of Error

A type I error occurs when a true null hypothesis is rejected. The probability of committing a type I error is called the level of significance of the test and is denoted by α. Hence

$$\alpha = P(\text{Type I error}) = P(\text{Reject} H_o | H_o \text{is true})$$

A type II error occurs when a false null hypothesis is not rejected. The probability of committing Type II error is denoted by β. Hence,

$$\begin{aligned} \beta &= P(\text{Type II error}) = P(\text{Fail to Reject} H_o | H_o \text{is false}) \\ \text{Power} &= 1 - \beta \end{aligned}$$

Hypothesis Testing using p-value

A p-value is the probability of observing a value of the test statistic as extreme or more extreme than the observed one, assuming the null hypothesis is true. This means the p-value is the smallest significance level at which the null hypothesis is rejected. For this reason, the p-value is sometimes called the observed significance level.

- If the computed p-value is smaller than α, reject the null hypothesis.
- If the computed p-value is greater than α, fail to reject the null hypothesis.

Remark: The p-value for a two sided hypothesis we simply double the tail probability of the test statistic. For example , suppose we are testing the following hypothesis:

$$H_o : \mu = 5$$
$$H_1 : \mu \neq 5$$

further suppose the computed test statistics is $z = 2.31$

Then the tail area associated with the test statistics is 0.0104. To compute the p-value, for a two sided hypothesis tests , we have P-value = 0.0104+0.0104=0.0208.

The p-value of 0.0208 is the likelihood of observing a value of the test statistic greater than 2.31 or less than -2.31 given the null hypothesis is true.

P-value

The p-value answers the following question:
If the null hypothesis is true, how likely is it that our observed test statistic takes the value we observed or more extreme? If this probability is small, then we reject the null hypothesis. If the p-value is not small, then we do not reject the null hypothesis.
Interpreting p-values.
Here are some rough guidelines for interpreting p-values which can be used in any testing scenario (not just for testing hypotheses about the mean). Let p denote the p-value of a test:

- If $p \leq 0.01$, then one has very strong evidence against the null hypothesis.
- If $0.01 < p \leq 0.05$, then one has strong evidence against the null hypothesis.
- If $0.05 < p < 0.10$, then the evidence against the null hypothesis is moderate to weak.
- If $0.10 \leq p < 0.20$ then the evidence against the null hypothesis is quite weak.

Hypothesis testing and significance testing on mean

There are three forms of test of hypotheses on the mean of a distribution:

Right Tailed test

$$H_0 : \mu \leq \mu_o, \qquad H_a : \mu > \mu_o$$

Left Tailed test

$$H_0 : \mu \geq \mu_o, \qquad H_a : \mu < \mu_o$$

Two Tailed test

$$H_0 : \mu = \mu_o, \qquad H_a : \mu \neq \mu_o$$

In order to test a hypothesis on a parameter , say μ, we must find a statistic whose probability distribution is known at least under the null hypothesis. This statistic is called the test statistic.

Hypothesis testing and significance testing on mean

If σ is known we have the test statistic

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

which follows standard normal distribution.
<u>Decision criteria:</u>

Hypothesis	Rejection Criteria
$H_o : \mu = \mu_o \quad Vs. \quad H_a : \mu > \mu_0$	$Z > Z_\alpha$
$H_o : \mu = \mu_o \quad Vs. \quad H_a : \mu < \mu_0$	$Z < -Z_\alpha$
$H_o : \mu = \mu_o \quad Vs. \quad H_a : \mu \neq \mu_0$	$\|Z\| > Z_{\alpha/2}$

P-value for Normal Distribution

The p-value is the smallest significance level at which the null hypothesis would still be rejected.
1. The p-value is the α that we can compute if we use the actual value of our test statistic as z.
2. For a z test (with normal or approximately normal populations) the p-value is

- $p = 1 - \Phi(z)$ for an upper tailed test
  p= 1-pnorm(z)
- $p = \Phi(z)$ for lower tailed test
  p= pnorm(z)
- $p = 2(1 - \Phi(|z|))$ for two tailed test
  $p = 2(1 - pnorm(|z|))$

Example

A new type of body armor is tested if it satisfies the specification of at most $\mu_0 = 1.9$ in of displacement when hit with a certain type of bullet. The manufacturer tests by firing one round each at 36 samples of the new armor and measuring the displacement upon impact. The result is a sample mean displacement of 1.91 in. Assume the displacements are normally distributed with mean μ and a standard deviation of 0.06 in. Test if the armor is up to specifications at the 10% significance level.

Solution: Parameter of interest is μ the true mean displacement.
Null hypothesis: $H_0 : \mu = 1.9$
Alternative hypothesis: $H_a : \mu > 1.9$
The test statistic is

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{1.91 - 1.9}{0.06/\sqrt{36}} = 1$$

Rejection region: We use an upper tailed test and the critical value is $Z_{0.1} = 1.2815$.
Decision: Fail to reject the null hypothesis as $1 < 1.2815$

Z-test using R

We need to install new package "TeachingDemos" or "BSDA"

```
z.test {TeachingDemos}  R Documentation
Z test for known population standard deviation

z.test(x,mu =0,stdev,alternative=c("two.sided","less","grea
 sd = stdev, conf.level = 0.95, ...)
```

Example:

```
x <- rnorm(25, 100, 5)
z.test(x, 99, 5)
```

Example

A new type of body armor is tested if it satisfies the specification of at most $\mu_0 = 1.9$ in of displacement when hit with a certain type of bullet. The manufacturer tests by firing one round each at 36 samples of the new armor and measuring the displacement upon impact. The result is a sample mean displacement of 1.91 in. Assume the displacements are normally distributed with mean μ and a standard deviation of 0.06 in. Test if the armor is up to specifications at the 10% significance level.

Solution: Parameter of interest is μ the true mean displacement.
Null hypothesis: $H_0 : \mu = 1.9$
Alternative hypothesis: $H_a : \mu > 1.9$

```
>library(BSDA)
>zsum.test(mean.x=1.91,sigma.x=0.06,n.x=36, mu=1.9,alt="greater")
        One-sample z-Test
data:  Summarized x
z = 1, p-value = 0.1587
alternative hypothesis: true mean is greater than 1.9
```

Decision: Fail to reject the null hypothesis as p-value $> \alpha$

Hypothesis testing and significance testing on mean

If σ is unknown and n is small we have the test statistic

$$T = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}$$

which follows $t-$ distribution with $n-1$ degrees of freedom.
Decision criteria:

Hypothesis	Rejection Criteria
$H_o : \mu = \mu_o$ Vs. $H_a : \mu > \mu_0$	$T > T_{\alpha,n-1}$
$H_o : \mu = \mu_o$ Vs. $H_a : \mu < \mu_0$	$T < -T_{\alpha,n-1}$
$H_o : \mu = \mu_o$ Vs. $H_a : \mu \neq \mu_0$	$\lvert T \rvert > T_{\alpha/2,n-1}$

P-value for Students's t- Distribution

The p-value is the smallest significance level at which the null hypothesis would still be rejected.
1. The p-value is the α that we can compute if we use the actual value of our test statistic as t.
2. For a t test (with normal or approximately normal populations) the p-value is

- $p = P(T > t)$ for an upper tailed test
  p= 1-pt(t, n)
- $p = P(T < t)$ for lower tailed test
  p= pt(t,n)
- $p = 2(P(T > t)$ for two tailed test
  $p = 2(1 - pt(|t|), n)$

Example

A certain medication is supposed to stay in the blood stream for at least 12 hours. A new pill design is tested in 25 patients. The sample average time the medication is detected at sufficient levels in the blood stream is 11.8 hours with a sample standard deviation of 0.5 hours. Does this data suggest that the actual mean time of sufficient levels in the blood stream is less than the desired 12 hours?

Solution: Parameter of interest: μ *the actual mean time of sufficient levels in the blood stream*

$$\begin{aligned} H_0 &: \quad \mu = 12 \\ Ha &: \quad \mu < 12 \end{aligned}$$

The test statistic is

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{11.8 - 12}{0.5/\sqrt{25}} = -2.0$$

p-value: Because the test is lower tailed: We have DF= 24,
$p = P(t \leq -0.2) = 0.0284$
Conclusion: At the 0.05 level, H_0 *would be rejected, at the* 0.01 *level it would not be rejected.*

Example

Suppose a car manufacturer claims a model gets 25 mpg. A consumer group asks 10 owners of this model to calculate their mpg and the mean value was 22 with a standard deviation of 1.5. Is the manufacturer's claim supported?

We want to test

$$H_0 : \mu = 25$$
$$H_a : \mu < 25$$

To test using R we simply need to tell R about the type of test. We need to calculate the test statistic and then find the p-value.

```
## Compute the t statistic.
> xbar=22; s=1.5;n=10
> t = (xbar-25)/(s/sqrt(n))
> t
[1] -6.324555
## use pt to get the distribution function of t
> pt(t,df=n-1)
[1] 6.846828e-05
```

A farmer want to test if a new brand of fertilizer increases his wheat yield per plot. He put the new fertilizer on 15 equal plots and records the yields for the 15 plots which are given below. If his traditional yield is two bushels per plot, conduct a test of significance for μ at $\alpha = 0.05$ significance level.

$$2.5, 3.0, 3.1, 4.0, 1.2, 5.0, 4.1, 3.9, 3.2, 3.3, 2.8, 4.1, 2.7, 2.9, 3.7$$

Solution: We would like to test

$$\begin{aligned} H_0 &: \mu = 2 \\ H_a &: \mu > 2 \end{aligned}$$

```
> x=c(2.5,3.0,3.1,4.0,1.2,5.0,4.1,3.9,3.2,3.3,2.8,4.1,2.7,2.9,3.
> t.test(x, alternative="greater", mu=2)
        One Sample t-test
data:  x
t = 5.6443, df = 14, p-value = 3.026e-05
alternative hypothesis: true mean is greater than 2
95 percent confidence interval:
 2.894334      Inf
```

Example

A random sample, consisting of the values listed below, was taken from a population which is normally distributed. Test the hypothesis that mean is 25 at $\alpha = 0.1$ and construct 90% confidence interval for population mean.

$$22, 23, 24, 22, 25, 26, 27, 25, 30, 26, 29$$

Solution: We would like to test

$$H_0 : \mu = 25 \qquad H_a : \mu \neq 25$$

```
 y=c(22, 23, 24, 22, 25, 26, 27, 25,30,26,29)
> t.test(y, mu=25)
        One Sample t-test
data:  y
t = 0.4607, df = 10, p-value = 0.6549
alternative hypothesis: true mean is not equal to 25
95 percent confidence interval:
 23.60476 27.12251
sample estimates:
mean of x
 25.36364
```

Comparing two population Means: Large Sample

Hypothesis testing on μ_1, μ_2:
Let a sample of size n_1 and n_2 from a population with means μ_1, μ_2 and variances σ_1^2, σ_2^2 yields sample means $\overline{x}_1, \overline{x}_2$ then the distribution of $\overline{x}_1 - \overline{x}_2$ has approximately normal distribution with

$$\mu_{\overline{x}_1 - \overline{x}_2} = \mu_1 - \mu_2$$

$$\sigma^2_{\overline{x}_1 - \overline{x}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

If the two samples are independent.
The test statistic is given by

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

and if $\sigma_1^2 = \sigma_2^2 = \sigma^2$ the test statistic is

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Comparing two population Means: Large Sample

Now, if we are interested in testing the hypothesis

$$H_o : \mu_1 = \mu_2$$
$$H_a : \mu_1 \neq \mu_2$$

we will reject the null hypothesis if $|z| > z_{\alpha/2}$.
Table below summarizes the rejection criteria.
Decision criteria:

Hypothesis	Rejection Criteria
$H_o : \mu_1 = \mu_2$ *Vs.* $H_1 : \mu_1 > \mu_2$	$z > z_\alpha$
$H_o : \mu_1 = \mu_2$ *Vs.* $H_1 : \mu_1 < \mu_2$	$z < -z_\alpha$
$H_o : \mu_1 = \mu_2$ *Vs.* $H_1 : \mu_1 \neq \mu_2$	$\lvert z \rvert > z_{\alpha/2}$

Also note that σ_1^2, σ_2^2 can be approximated by s_1^2 and s_2^2 if both n_1, n_2 are large.

Comparing two population Means: Small Sample

Assumptions for a small sample test of $\mu_1 - \mu_2$
(i) Both populations of interest are normally distributed.
(ii) Both the populations have approximately equal(but unknown) variance
In order to test

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2 \end{aligned}$$

We consider the test statistic

$$t_0 = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ is called pooled variance. Also s_1^2 and s_2^2 are the variances of the first and second samples.
Decision criteria:

Hypothesis	Rejection Criteria
$H_0 : \mu_1 = \mu_2 \quad Vs. \quad H_1 : \mu_1 > \mu_2$	$t_0 > t_{\alpha, n_1+n_2-2}$
$H_0 : \mu_1 = \mu_2 \quad Vs. \quad H_1 : \mu_1 < \mu_2$	$t_0 < -t_{\alpha, n_1+n_2-2}$
$H_0 : \mu_1 = \mu_2 \quad Vs. \quad H_1 : \mu_1 \neq \mu_2$	$\lvert t_0 \rvert > t_{\alpha/2, n_1+n_2-2}$

Comparing two population Means: Small Sample

Assumptions for a small sample test of $\mu_1 - \mu_2$
(i) Both populations of interest are normally distributed.
(ii) The population variances are unequal
In order to test

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2 \end{aligned}$$

We consider the test statistic

$$t_0 = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

Also s_1^2 and s_2^2 are the variances of the first and second samples. Note that t_0 has student's t distribution with degrees of freedom given by

$$\nu = \left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 \times \left(\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}\right)^{-1}$$

Decision criteria:

Hypothesis	Rejection Criteria
$H_0 : \mu_1 = \mu_2$ *Vs.* $H_1 : \mu_1 > \mu_2$	$t_0 > t_{\alpha,\nu}$
$H_0 : \mu_1 = \mu_2$ *Vs.* $H_1 : \mu_1 < \mu_2$	$t_0 < -t_{\alpha,\nu}$
$H_0 : \mu_1 = \mu_2$ *Vs.* $H_1 : \mu_1 \neq \mu_2$	$\lvert t_0 \rvert > t_{\alpha/2,\nu}$

Inferences about the differences in Means, Paired t-test

To test

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

Let $\mu_d = \mu_1 - \mu_2$ then the above testing problem is equivalent to

$$H_0 : \mu_d = 0$$
$$H_1 : \mu_d \neq 0$$

The test statistic for this hypothesis is

$$t_0 = \frac{\overline{d}}{S_d/\sqrt{n}}$$

where $\overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ is the mean of the differences and S_d is the sample standard deviation of the differences.

Decision criteria:

Hypothesis			Rejection Criteria
$H_0 : \mu_1 = \mu_2$	$Vs.$	$H_1 : \mu_1 > \mu_2$	$t_0 > t_{\alpha,n-1}$
$H_0 : \mu_1 = \mu_2$	$Vs.$	$H_1 : \mu_1 < \mu_2$	$t_0 < -t_{\alpha,n-1}$
$H_0 : \mu_1 = \mu_2$	$Vs.$	$H_1 : \mu_1 \neq \mu_2$	$\lvert t_0 \rvert > t_{\alpha/2,n-1}$

Example- Two sample Z test

Freshman at public universities work 12.2 hours per week for pay on the average, with a standard deviation of 10.5 hours. At private universities, the average for freshman is 10.2 hours, with a standard deviation of 9.9 hours. The sample size for each is 1,000. Is the difference between the averages real or is it just chance variation. Perform a level 0.05 independent two-sample test to find out.

Solution: Let Let μ_1 *and* μ_2 *denote the average number of hours the public and private university freshman students work per week respectively. We want test*

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_a &: \mu_1 \neq \mu_2 \end{aligned}$$

We have $n_1 = 1000, n_2 = 1000, \bar{x}_1 = 12.2, \bar{x}_2 = 10.2, \sigma_1 = 10.5, \sigma_2 = 9.9$ *We have the test statistic*

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(12.2 - 10.2) - 0}{\sqrt{\frac{10.5^2}{1000} + \frac{9.9^2}{1000}}} = 4.38$$

`p-value=2*pnorm(-4.38)=0.0000118`

Decision: Reject the null hypothesis.

Example- Two sample Z test

Freshman at public universities work 12.2 hours per week for pay on the average, with a standard deviation of 10.5 hours. At private universities, the average for freshman is 10.2 hours, with a standard deviation of 9.9 hours. The sample size for each is 1,000. Is the difference between the averages real or is it just chance variation. Perform a level 0.05 independent two-sample test to find out.

Solution: Let Let μ_1 and μ_2 denote the average number of hours the public and private university freshman students work per week respectively. We want test

$$
\begin{aligned}
H_0 &: \mu_1 = \mu_2 \\
H_a &: \mu_1 \neq \mu_2
\end{aligned}
$$

```
> library(BSDA)
> zsum.test(mean.x=12.2,sigma.x=10.5,n.x=1000,
     mean.y=10.2, sigma.y=9.9, n.y=1000)
      Two-sample z-Test
data:  Summarized x and y
z = 4.3826, p-value = 1.173e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.10556 2.89444
```

Example- Unknown but Equal Variances

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:

```
Drug :      15 10 13 7  9 8 21  9 14  8
Placebo:    15 14 12 8 14 7 16 10 15 12
```

Suppose the assumptions of equal variances and normality are valid. Perform the test to show that the drug group has smaller mean.

Let μ_1 and μ_2 respectively denote the mean of drug group and placebo group. We want to test the following

$$\begin{aligned} H_0 &: \mu_1 \geq \mu_2 \\ H_a &: \mu_1 < \mu_2 \end{aligned}$$

Example

We can use R to test the hypothesis

```
> x = c(15, 10, 13, 7, 9, 8, 21, 9, 14, 8)
> y = c(15, 14, 12, 8, 14, 7, 16, 10, 15, 12)
> t.test(x,y,alt="less",var.equal=TRUE)
        Two Sample t-test
data:  x and y
t = -0.5331, df = 18, p-value = 0.3002
alternative hypothesis: true difference in means is less th
95 percent confidence interval:
     -Inf 2.027436
sample estimates:
mean of x mean of y
     11.4      12.3
```

Decision: Not enough evidence to reject the null hypothesis.

Example: Paired t-test

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 the mileage for that tank was recorded. The mileage was recorded again for the same cars using the other kind of gasoline.

```
> Premium = c(19, 22, 24, 24, 25, 25, 26, 26, 28, 32)
> Regular = c(16, 20, 21, 22, 23, 22, 27, 25, 27, 28)
> t.test(Premium,Regular, alternative="greater", paired=TRUE)
        Paired t-test
data:  Premium and Regular
t = 4.4721, df = 9, p-value = 0.0007749
alternative hypothesis: true difference in means is greater than
95 percent confidence interval:
 1.180207      Inf
sample estimates:
mean of the differences
                      2
```

Decision: Reject the null hypothesis. There is strong evidence of a mean

Example

Do the volume measurements of tumor change based on physician?
The volume of the tumor was measured by two separate physicians under similar condition. Data below are recorded.
Dr.1 :15.8, 22.3, 14.5, 15.7, 26.8, 24.0, 21.8, 23.0, 29.3, 20.5
Dr.2 :17.2, 20.3, 14.2, 18.5, 28.0, 24.8, 20.3, 25.4, 27.5, 19.7

```
> Dr.1=c(15.8, 22.3, 14.5, 15.7, 26.8, 24.0, 21.8, 23.0, 29.3, 20.5)
> Dr.2=c(17.2, 20.3, 14.2, 18.5, 28.0, 24.8, 20.3, 25.4, 27.5, 19.7)
> t.test(Dr.1,Dr.2, paired=T)

        Paired t-test
data:  Dr.1 and Dr.2
t = -0.3989, df = 9, p-value = 0.6993
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.467632  1.027632
sample estimates:
mean of the differences
                  -0.22
```

Decision: No enough evidence that the volume measurements of tumor change based on physician

Example: Two sample t Vs. Paired t

The composite biodiversity score based on a kick sample of aquatic invertebrates are provided below.

Down: 20 15 10 5 20 15 10 5 20 15 10 5 20 15 10 5

Up: 23 16 10 4 22 15 12 7 21 16 11 5 22 14 10 6

The elements are paired because the two samples were taken on the same river, one upstream and one downstream from the same sewage outfall.

If we ignore the fact that they are paired we may get completely different results

```
> down=c(20,15,10,5,20,15,10,5,20,15,10,5,20,15,10,5)
> up=c(23,16,10,4,22,15,12,7,21,16,11,5,22,14,10,6)
> t.test(down,up)
        Welch Two Sample t-test
data:  down and up
t = -0.4088, df = 29.755, p-value = 0.6856
alternative hypothesis: true difference in means is not equal to
95 percent confidence interval:
 -5.248256  3.498256
```

Example: Two sample t Vs. Paired t

```
> down=c(20,15,10,5,20,15,10,5,20,15,10,5,20,15,10,5)
> up=c(23,16,10,4,22,15,12,7,21,16,11,5,22,14,10,6)
> t.test(down,up,paired=T)

        Paired t-test
data:  down and up
t = -3.0502, df = 15, p-value = 0.0081
alternative hypothesis: true difference in means is not equal to
95 percent confidence interval:
 -1.4864388 -0.2635612
sample estimates:
mean of the differences
                 -0.875
```

Example

A "new diet and exercise" program has been advertised to be a remarkable way to reduce blood glucose levels in diabetic patients. Ten randomly selected diabetic patients are put on the program and here are the results after one month are given by the following table:

Before:	268	225	252	192	307	228	246	298	231	185
After:	106	186	223	110	203	101	211	176	194	203

Do the data provide sufficient evidence to the claim that the new program reduces blood glucose level in diabetic patients? Use $a = 0.05$.

```
> Before=c( 268, 225, 252, 192, 307, 228, 246, 298, 231, 185)
> After=c( 106, 186, 223, 110, 203,  101, 211, 176, 194, 203)
> t.test(Before, After, alternative="greater", paired=T)
        Paired t-test
data:  Before and After
t = 4.0489, df = 9, p-value = 0.001445
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 39.34775      Inf
sample estimates: mean of the differences
                   71.9
```

Decision: The new diet and exercise program is effective.