

Lab 5

!!Please read C8 and C9.

One sample t-test

-for the mean of a normal distribution with unknown variance.

Two-tailed test:

$H_0 : \mu = 5$ with the alternative $H_A : \mu \neq 5$

```
x<-rnorm(100,2,2)
# we do not test for normality of the data as x is generated normal
t.test(x,alternative="two.sided",mu=5)
```

One Sample t-test

```
data: x
t = -15.193, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 5
95 percent confidence interval:
 1.443943 2.265475
sample estimates:
mean of x
1.854709
```

H_0 is rejected as the $p\text{-value} < 0.01$ (or equivalent $t < -qt(0.99, 99)$) – we can say that the data do not have the mean=5 with a confidence level of 0.99.

Obs. When you run the code, the R's response will be slightly different as x is randomly generated.

The meaning of the outcomes of the test in R:

t is the test statistic ($=\sqrt{100} * (\text{mean}(x) - 5) / \sqrt{\text{var}(x)}$) – see C9 slide 4)

df is the number of degrees of freedom ($=100-1$)

the p-value is computed based on the formula in C9 slide 8

95 percent confidence interval:

1.443943 2.265475

is the interval that has 0.95 probability of containing the true value of the mean

This interval is calculated as follows:

We know that $\frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{n-1}$ (C8 slide 21), hence $P\left(-t_{n-1; 1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{S} \leq t_{n-1; 1-\frac{\alpha}{2}}\right) = 1 - \alpha$ (take $\alpha=0.05$).

Rewrite the inequalities $-t_{n-1; 1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\bar{X} - \mu)}{S} \leq t_{n-1; 1-\frac{\alpha}{2}}$ as $lower_limit \leq \mu \leq upper_limit$

and thus, we find the 95% confidence interval for the mean.

The formulas in R are:

`[-qt(0.975,99)*sqrt(var(x))/sqrt(100)+mean(x) , qt(0.975,99)*sqrt(var(x))/sqrt(100)+mean(x)]`

The operating characteristic function (C8 slides 25,26) is:

```
oc<-function(x,mu,mu0,alpha){
  n<-length(x)
  s<-sqrt(var(x))
  return (pt(qt(1-alpha/2, df = n-1)-sqrt(n)*(mu-mu0)/s, df = n-1)
          -pt(-qt(1-alpha/2, df = n-1)-sqrt(n)*(mu-mu0)/s, df = n-1) )
}

#1. keep the sample size n constant
mu=2      #the true value of the parameter
mu0=seq(2.1,by=0.1,3)
plot(mu0-mu,oc(x,mu,mu0,0.05))           #type II error

#2. keep mu and mu0 constant and vary the sample size n
mu=2      #the true value of the parameter
mu0=2.3

r<-oc(x,mu,mu0,0.05)

for (n in seq(200,by=100,1000))
  r<-c(r, oc(rnorm(n,2,2),mu,mu0,0.05))

plot(r)
```

What do these graphs tell us about the probability of type II error? (see C8 slides 18,19)

Left-tailed test:

$H_0 : \mu = 5$ with the alternative $H_A : \mu < 5$

```
t.test(x,alternative="less",mu=5)
```

One Sample t-test

```
data: x
t = -15.193, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is less than 5
95 percent confidence interval:
 -Inf 2.198438
sample estimates:
mean of x
1.854709
```

Right-tailed test:

$H_0 : \mu = 5$ with the alternative $H_A : \mu > 5$

```
t.test(x,alternative="greater",mu=5)
```

One Sample t-test

```
data: x
t = -15.193, df = 99, p-value = 1
alternative hypothesis: true mean is greater than 5
95 percent confidence interval:
 1.510981      Inf
sample estimates:
mean of x
1.854709
```

How do you interpret these results?

```
#recall that x is rnorm(100,2,2)
```

```
t.test(x,alternative="less",mu=2.23)
```

```
t.test(x,alternative="two.sided",mu=2.23)
```

Look at the outcomes of these two tests above (after multiple generations of x). What do you notice? Draw the same plot as in C9 slide 5.

The following comments are taken from <https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-the-differences-between-one-tailed-and-two-tailed-tests/>

“When is a one-tailed test appropriate?”

Because the one-tailed test provides more power to detect an effect, you may be tempted to use a one-tailed test whenever you have a hypothesis about the direction of an effect. Before doing so, consider the consequences of missing an effect in the other direction. Imagine you have developed a new drug that you believe is an improvement over an existing drug. You wish to maximize your ability to detect the improvement, so you opt for a one-tailed test. In doing so, you fail to test for the possibility that the new drug is less effective than the existing drug. The consequences in this example are extreme, but they illustrate a danger of inappropriate use of a one-tailed test.

So when is a one-tailed test appropriate? If you consider the consequences of missing an effect in the untested direction and conclude that they are negligible and in no way irresponsible or unethical, then you can proceed with a one-tailed test. For example, imagine again that you have developed a new drug. It is cheaper than the existing drug and, you believe, no less effective. In testing this drug, you are only interested in testing if it less effective than the existing drug. You do not care if it is significantly more effective. You only wish to show that it is not less effective. In this scenario, a one-tailed test would be appropriate.

When is a one-tailed test NOT appropriate?

Choosing a one-tailed test for the sole purpose of attaining significance is not appropriate. Choosing a one-tailed test after running a two-tailed test that failed to reject the null hypothesis is not appropriate, no matter how "close" to significant the two-tailed test was. Using statistical tests inappropriately can lead to invalid results that are not replicable and highly questionable—a steep price to pay for a significance star in your results table!”

Two sample t-test (non-paired)

The “sleep” dataset contains data which show the effect of two soporific drugs (increase in hours of sleep) on groups consisting of 10 patients each.

```
sleep
group1<- sleep[sleep['group']=='1', 'extra']
group2<- sleep[sleep['group']=='2', 'extra']
```

First, check the normality of the data (use the Shapiro-Wilk Normality Test for small sized samples, $n \geq 3$)

```
shapiro.test(group1)
```

Shapiro-Wilk normality test

```
data: group1
W = 0.92581, p-value = 0.4079
```

```
shapiro.test(group2)
```

Shapiro-Wilk normality test

```
data: group2
W = 0.9193, p-value = 0.3511
```

As both p-values are > 0.05 , we accept (fail to reject) the null hypotheses of normality.

Now, we use the F test to compare the variances of the two groups:

```
var.test(group1,group2)
```

F test to compare two variances

```
data: group1 and group2
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
```

We fail to reject the null hypothesis of equal variances.

We compare the means of the two groups:

```
t.test(group1,group2,alternative="two.sided",mu=0,paired=FALSE,var.equal=TRUE)
```

Two Sample t-test

```
data: group1 and group2
t = -1.8608, df = 18, p-value = 0.07919
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-3.363874  0.203874
sample estimates:
mean of x mean of y
 0.75      2.33
```

As the $p\text{-value} > 0.05$, the conclusion is that the two drugs do not have significantly different effects on the tested patients.

Apply now the two sample t-test (non-paired) for the data

```
X<-rnorm(34,4,3.2)
```

```
Y<-rnorm(26,4.5,1)
```

Pay attention to the parameters of t.test(...).
What is the conclusion of the test?

Two sample t-test (paired)

(X; Y) = (blood pressure before treatment; blood pressure after treatment)

We generate data for 50 patients

```
mu<-c(16,14)
sigma<-matrix(c(5,1.5,1.5,1),2,2)
library(MASS)
Z<-mvrnorm(n = 50, mu, sigma, empirical = FALSE)
X<-c(Z[,1])
Y<-c(Z[,2])
```

```
#the test for normality can be done for X-Y using shapiro.test(X-Y) -- not
#necessary here as the data is generated as multivariate normal
```

```
t.test(X, Y, paired = TRUE)
```

Paired t-test

data: X and Y

t = 8.5318, df = 49, p-value = 2.972e-11

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

1.667491 2.695042

sample estimates:

mean of the differences

2.181267

We reject the null hypothesis of equal means (i.e. there are significant differences before/after the treatment).

Apply the paired two-sample t-test for the data below (taken from <https://bolt.mph.ufl.edu/6050-6052/unit-4b/module-13/paired-t-test/#moredata>)

```
a<-c(1903,1935,1910,2496,2108,1961,2060,1444,1612,1316,1511)
b<-c(2009,1915,2011,2463,2180,1925,2122,1482,1542,1443,1535)
```