# R Tutorial for STAT 350 Lab 4

**Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi**

### 1. Confidence Interval for One Population

To have R calculate a confidence interval, we use the "t.test()" command. As can be seen by name this uses the t test, not the z test. I will explain in Chapter 7 why this is the better methodology. In this lab, we will just be showing you which part of the output is necessary for this Lab; the complete output will be explained in Lab 6.

**Example (DATA SET: DMS.txt – website)** Many food products contain small quantities of substances that would give an undesirable taste or smell if they were present in large amounts. An example is the "off-odors" caused by sulfur compounds in wine. Oenologists (wine experts) have determined the odor threshold, the lowest concentration of a compound that the human nose can detect. For example, the odor threshold for dimethyl sulfide (DMS) is given in the oenology literature as 25 micrograms per liter of wine ($\mu$g/l). Untrained noses may be less sensitive, however. Here are the DMS odor thresholds for 10 beginning students of oenology:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 31 | 31 | 43 | 36 | 23 | 34 | 32 | 30 | 20 | 24 |

(a) Generate a 95% confidence interval for the mean DMS odor threshold among all beginning oenology students.

**Solution:**

```
>wine=read.table(file="DMS.txt",header=T)
>wine

>t.test(wine$DMS,conf.level=0.95)

        One Sample t-test

data:  wine$DMS
t = 14.2361, df = 9, p-value = 1.775e-07
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 25.56935 35.23065
sample estimates:
mean of x
     30.4
```

The 95% confidence interval is (25.56935, 35.23065)

### 2. Calculating the power and generating a power curve

# R Tutorial for STAT 350 Lab 4

**Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi**

To calculate the power, we want to know the probability that we will reject $H_0$ when it is false. To perform power calculations, you need to assume the real value of μ. Therefore, if $H_a$: $\mu > \mu_0$, we want to calculate $P(\bar{X} \geq x^* \mid \mu = \mu^*, \mu^* > \mu_0)$ where $\mu^*$ is a specified value greater than $\mu_0$ and $x^*$ is the first value x, such that (s.t.) $x^* > x$ => Reject $H_0$: $\mu = \mu_0$.

**Example** Consider the following competing hypotheses:

$H_0$: $\mu = 500$

$H_a$: $\mu \neq 500$

at the 1% level of significance assuming the population standard deviation, \sigma, is known.

(a) Generate a power curve.
(b) What sample size would be required for the power to be at least 0.80 against the specific alternative of μ = 498.5?

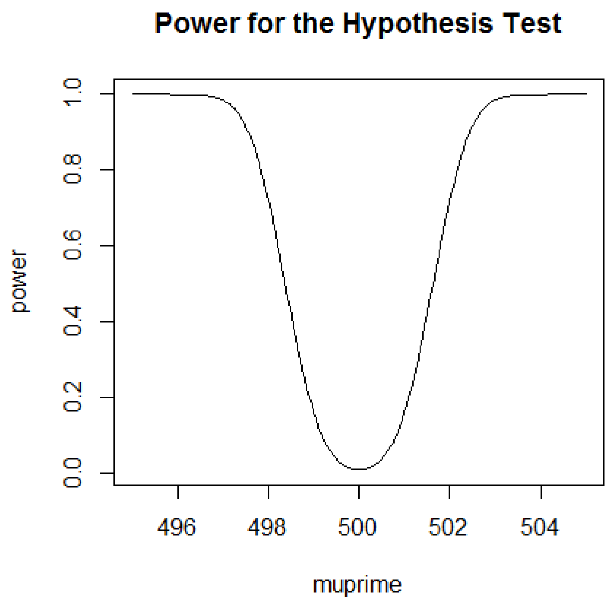
### Solution:

#### (a) Generate a power curve.

```
# variables to change:
#  n, alpha, mu0, sigma
#  muprime: the real values of mu
#    by 0.05 - this needs to be fine enough so that the curve
#               looks smooth

n <- 10
alpha <- 0.01
mu0 <- 500
sigma <- 2
sigman <- sigma/sqrt(n) #standard error
#z is from alpha/2 for a 2-tailed test
z <- qnorm(1 - alpha/2)
muprime <- seq(from=495, to=505, by=0.05)
x1 <- mu0 - z*sigman #Value for which x < x1 => H_0 will be rejected
x2 <- mu0 + z*sigman #Value for which x > x2 => H_0 will be rejected
px1 <- pnorm(x1, muprime, sigman) #CDF up to x1 for various muprime
px2 <- pnorm(x2, muprime, sigman, lower.tail = FALSE) #S_X(x2): muprime
power <- px1 + px2 #Left Tail + Right Tail (Tails NOT symmetric!)
beta <- 1 - power #When obtaining by 'hand' beta is easier to find first

power

plot(muprime,power,main="Power for the Hypothesis Test",type="l")
```

# R Tutorial for STAT 350 Lab 4

**Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi**




Note: To answer the questions in Part 1; just change put in the appropriate value in the following command, that is replace ‘seq’ with the appropriate value.

```
muprime <- seq (from=495,to=505, by=0.05)
```

**(b) What sample size would be required for the power to be at least 0.80 against the specific alternative of μ = 498.5?**

```
#this time, I will fix muprime and change n;
n <- 10:100
muprime <- 498.5
alpha <- .01
mu0 <- 500
sigma <- 2
sigman <- sigma/sqrt(n)
z <- qnorm(1 - alpha/2)
x1 <- mu0 - z*sigman #Value for which x < x1 => H_0 will be rejected
x2 <- mu0 + z*sigman #Value for which x > x2 => H_0 will be rejected
px1 <- pnorm(x1, muprime, sigman) #CDF up to x1 for various muprime
px2 <- pnorm(x2, muprime, sigman, lower.tail = FALSE) #S_X(x2): muprime
power <- px1 + px2 #Left Tail + Right Tail (Tails NOT symmetric!)
beta <- 1 - power #When obtaining by ‘hand’ beta is easier to find first

answer <- data.frame(n, power)
answer
```

To determine the answer, look at the print out for the value of n that will produce the power that is desired.

**Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi**


```
    n     power
1   10 0.4191298
…
11  20 0.7817959
12  21 0.8054092
13  22 0.8268992
14  23 0.8463833
…
```

I truncated the output to only the relevant parts. You may delete all lines except for the title line and the line that contains the n that you are interested in.

Therefore, n = 21 is the first value which has a power greater than 0.8.

This could also be visualized with the following lattice graphics code:

```
library(lattice)
xyplot(power ~ n,
      panel = function(x, y){
            panel.xyplot(x, y)
            panel.abline(h = 0.8)
            panel.abline(v = 20)
      })
```