

```
#####
# L3.1. HYPOTHESES ABOUT MEAN OF A POPULATION
#####
## clear memory
rm(list = ls())
##-----
## L3.1.1. Hypotheses about mean
##-----
##Number of living cells in 5 wells under some conditions are given below.
##In a reference literature source authors claimed a mean quantity of 5000
##living cells under the same conditions. Is our result significantly different?

## put number of cells
x = c(5128, 4806, 5037, 4231, 4222)

## make analysis by hands (just to try)
n=length(x)
m=mean(x)
s=sd(x)
mu=5000
t=(m-mu)/s*sqrt(n)
p.val.1 = pt(t, df=n-1)
p.val.2 = 2*pt(t, df=n-1)

## make analysis t.test()
t.test(x, mu=5000)

##-----
## L3.1.2. Hypotheses about proportion
##-----
##During a study of a new drug against viral infection, you have found that 70
##out of 100 mice survived, whereas the survival after the standard therapy is
##60% of the infected population. Is this enhancement statistically significant?
p = 0.7
p0 = 0.6
n=100
## make analysis by hands
sp = sqrt(p0*(1-p0)/n)
z = (p-p0)/sp
p.val.1 = 1-pnorm(z)
p.val.1 ## Approximate!

## make analysis by prop.test()
prop.test(x=70, n=100, p=0.6, alternative="greater") ## Approximate!
## make analysis by binom.test()
binom.test(x=70, n=100, p=0.6, alternative="greater") ## Exact!

#####
# L3.2. HYPOTHESES ABOUT MEANS OF TWO POPULATION
#####
##-----
## L3.2.1. Unmatched samples
##-----
Mice=read.table("http://edu.sablab.net/data/txt/mice.txt", header=T, sep="\t")
str(Mice)
```

```

## calculate p-value by hands
y=Mice$Weight.change[Mice$Sex=="m"]
x=Mice$Weight.change[Mice$Sex=="f"]
s1=sd(x)^2/length(x)
s2=sd(y)^2/length(y)
t=(mean(x)-mean(y))/sqrt(s1+s2)
df=(s1+s2)^2/(s1^2/(length(x)-1)+s2^2/(length(y)-1))
2*pt(-abs(t),df)

## calculate p-value by t.test
t.test(x,y)

##-----
## Let us plot the distributions in a nice way
x11()
par(mfrow=c(3,2))
boxplot(Ending.weight~Sex,data=Mice,col=c("pink","lightblue"),outline=F)
title("Ending weight (g)")
plot(density(Mice$Ending.weight[Mice$Sex=="m"]),col="blue",lwd=2,main="")
lines(density(Mice$Ending.weight[Mice$Sex=="f"]),col="red",lwd=2)
title("Ending weight distributions")

## The distributions of weight change
boxplot(Weight.change~Sex,data=Mice,col=c("pink","lightblue"),outline=F)
title("Weights change")
plot(density(Mice$Weight.change[Mice$Sex=="m"]),col="blue",lwd=2,main="")
lines(density(Mice$Weight.change[Mice$Sex=="f"]),col="red",lwd=2)
title("Distributions of weight change")

## The distributions of bleeding time
boxplot(Bleeding.time~Sex,data=Mice,col=c("pink","lightblue"),outline=F)
title("Bleeding time")
plot(density(Mice$Bleeding.time[Mice$Sex=="m"],na.rm=T),col="blue",lwd=2,main="")
lines(density(Mice$Bleeding.time[Mice$Sex=="f"],na.rm=T),col="red",lwd=2)
title("Distributions of bleeding time")

## Perform t.tests (by default - unpaired, two.sided)
xm = Mice$Ending.weight[Mice$Sex=="m"]
xf = Mice$Ending.weight[Mice$Sex=="f"]
t.test(xm,xf)
wilcox.test(xm,xf)

xm = Mice$Weight.change[Mice$Sex=="m"]
xf = Mice$Weight.change[Mice$Sex=="f"]
t.test(xm,xf)
wilcox.test(xm,xf)

xm = Mice$Bleeding.time[Mice$Sex=="m"]
xf = Mice$Bleeding.time[Mice$Sex=="f"]
t.test(xm,xf)
wilcox.test(xm,xf)

## Compare the results!

##-----
## L3.2.2. Matched samples
##-----

```

```

BP=read.table("http://edu.sablab.net/data/txt/bloodpressure.txt", header=T, sep="\t")
str(BP)

## Unpaired
t.test(BP$BP.before, BP$BP.after)
wilcox.test(BP$BP.before, BP$BP.after)

## Paired
t.test(BP$BP.before, BP$BP.after, paired=T)
wilcox.test(BP$BP.before, BP$BP.after, paired=T)

##-----
## L3.2.3. Proportions
##-----
Mice=read.table("http://edu.sablab.net/data/txt/mice.txt", header=T, sep="\t")
str(Mice)

x = Mice$Sex[Mice$Strain == "SWR/J"]
y = Mice$Sex[Mice$Strain == "MA/MyJ"]

prop.test(c(sum(x=="m"), sum(y=="m")), n=c(length(x), length(y)))
## try with the same numbers, but with continuity correction
prop.test(c(9, 15), n=c(19, 23), correct=F)

## Perform tests (by default - unpaired, two.sided)
## t-test and Wilcoxon Rank Sum (equiv. Mann-Whitney test)

## Ending.weight
xm = Mice$Ending.weight[Mice$Sex=="m"]
xf = Mice$Ending.weight[Mice$Sex=="f"]
t.test(xm, xf)
wilcox.test(xm, xf)

## Weight.change
xm = Mice$Weight.change[Mice$Sex=="m"]
xf = Mice$Weight.change[Mice$Sex=="f"]
t.test(xm, xf)
wilcox.test(xm, xf)

## Bleeding.time
xm = Mice$Bleeding.time[Mice$Sex=="m"]
xf = Mice$Bleeding.time[Mice$Sex=="f"]
t.test(xm, xf)
wilcox.test(xm, xf)

#####
# L3.3. TESTING HYPOTHESIS ABOUT VARIANCES
#####
Bus=read.table("http://edu.sablab.net/data/txt/schoolbus.txt", header=T, sep="\t")
str(Bus)

## apply F-test
var.test(Bus[, 1], Bus[, 2])

```

```
#####
# L3.4. TESTING HYPOTHESIS ABOUT CORRELATION
#####

##-----
## L3.4.1. Significance of correlation and confidence interval
##-----

Chiton = read.table("http://edu.sablab.net/data/txt/chiton.txt", header=T, sep="\t")
str(Chiton)

cor.test(Chiton$Length, Chiton$Width)

##-----
## L3.4.2. Compare correlations
##-----

## dataset 1
x1 = runif(10)
y1 = 2*x1+rnorm(10, 0, 0.25)
r1 = cor(x1, y1)
## dataset 2
x2 = runif(20)
y2 = 2*x2+rnorm(20, 0, 0.5)
r2 = cor(x2, y2)

## define function - Fisher's transformation for correlation
fisher.r2z = function(r) { 0.5 * (log(1+r) - log(1-r)) }

Z=(fisher.r2z(r1)-fisher.r2z(r2))/sqrt((length(x1)-3)^(-1)+(length(x2)-3)^(-1))
## 2-tail pval
print(2*pnorm(-abs(Z)))

#####
# L3.5. Power analysis
#####
## calculate power for given n
power.t.test(n = 20, delta = 1)
## calculate n for given power
power.t.test(power = .90, delta = 1)

library(pwr)
?pwr.t.test
```