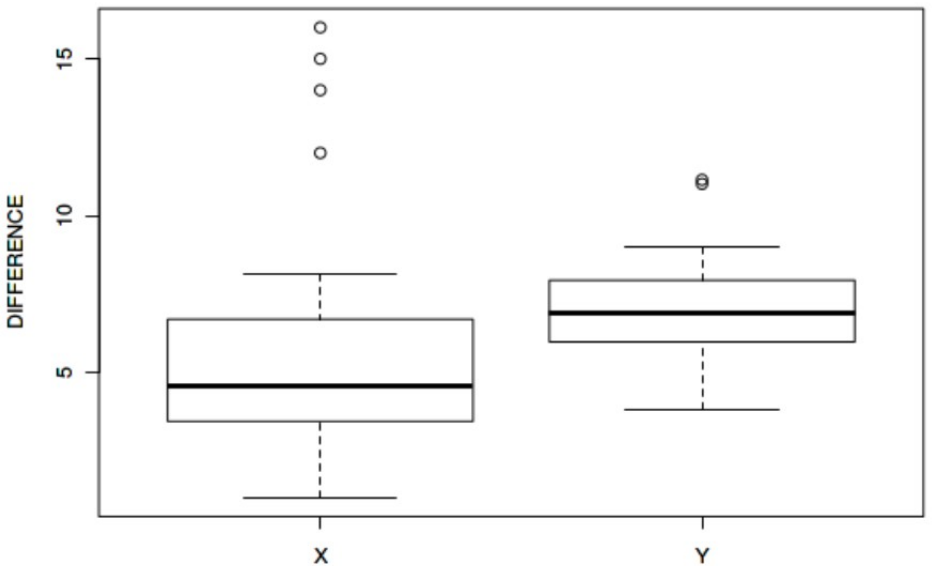
Two-Groups Statistical Test

```
set.seed(12)
X<-c(rnorm(20,mean=5,sd=2),12,15,14,16)
Y<-c(rnorm(24,mean=7,sd=2))
boxplot(X,Y,ylab="DIFFERENCE",names=c("X","Y"))
```



Parametric Statistical Test Fails

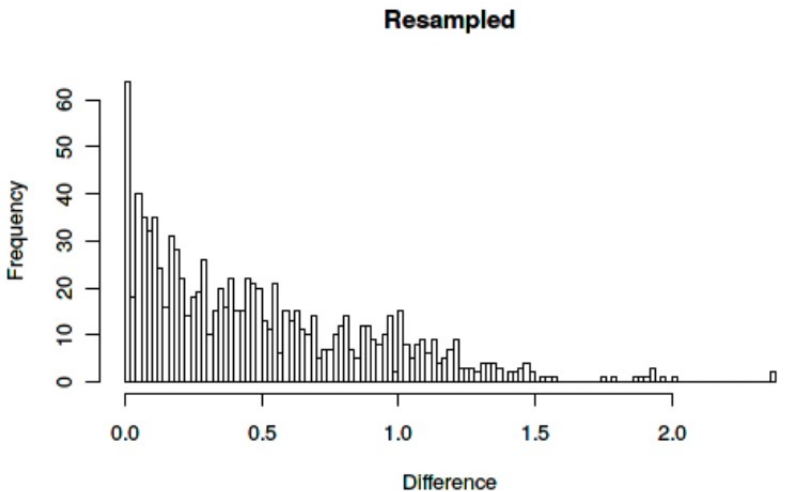

```
t.test(X,Y)
```

```
##
##     Welch Two Sample t-test
##
## data:  X and Y
## t = -1.084, df = 31.678, p-value = 0.2865
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2.8819545  0.8804344
## sample estimates:
## mean of x mean of y
##  5.989652  6.990412
```

Resampling

```
observed <- median(Y)-median(X)
print(paste0("Observed difference = ",observed))
res <- vector(length=1000)
for(i in 1:1000){Z<-c(X,Y); Y_res<-sample(Z,length(Y),FALSE);
X_res<-sample(Z,length(X),FALSE); res[i]<-median(Y_res)-median(X_res)}
hist(abs(res), breaks=100, main="Resampled", xlab="Difference")
print(paste0("p-value = ", sum(abs(res) >= abs(observed))/1000))
```

```
## [1] "Observed difference = 2.33059085402647"
## [1] "p-value = 0.002"
```



Compare with Mann-Whittney U test:

wilcox.test(X, Y)

Practice parametric vs. non-parametric correlation:

```
set.seed(12)

n<-10

x<-rnorm(n)
y<-3+2*x+10*rnorm(n)
plot(y~x)

x<-c(x,3)
y<-c(y,30)

plot(y~x)

cor.test(x,y)
cor.test(x,y,method="spearman")
```