STAT 2060 - BENJAMIN BUSS - Inference about two population means

3 Possible Approaches

- Classical(not used in statistical software[why?])
- P-value
- Confidence Interval

If population standard deviation is known used normal distribution, otherwise use t(assuming it fits)

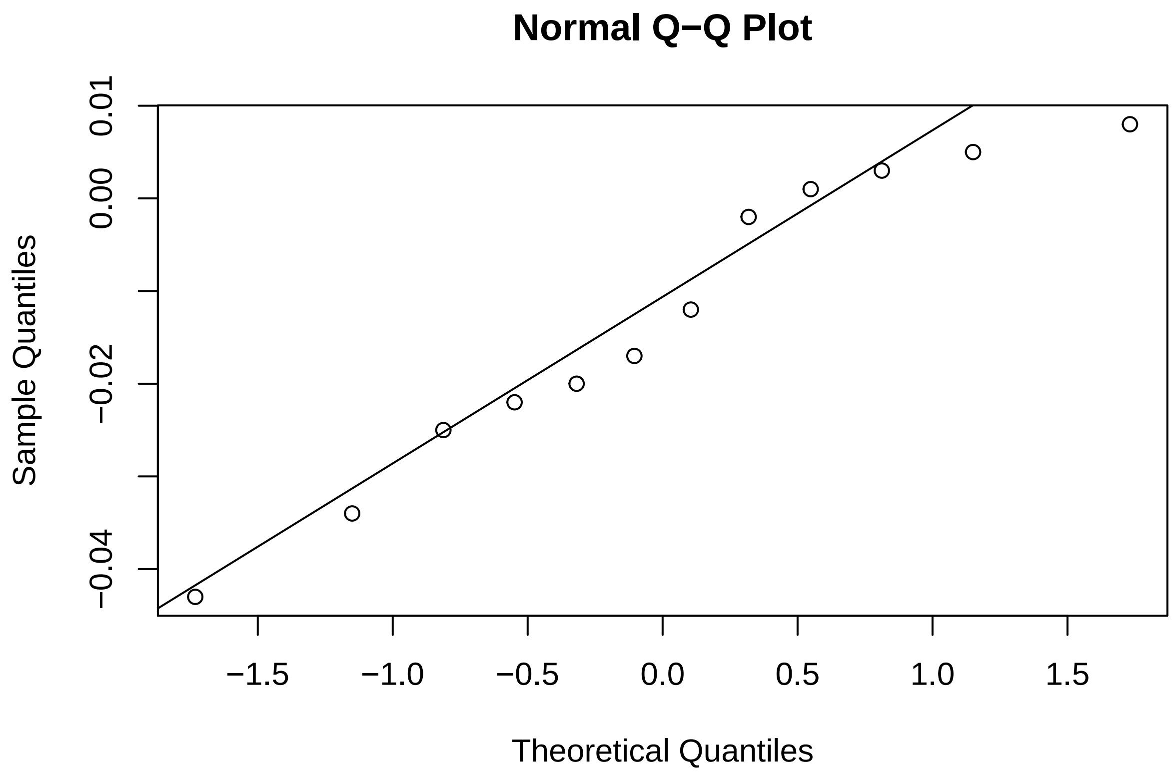
```
hands <- read.csv("C:/Users/benja/Downloads/hands.csv")
# Because the datasets are dependant, we can reduce the two variables onto one, the difference
diff <- hands$Dominant.Hand - hands$Nondominant.Hand; diff
```

```
##  [1] -0.002  0.008 -0.022  0.005 -0.017  0.001 -0.034  0.003 -0.043 -0.012
## [11] -0.020 -0.025
```

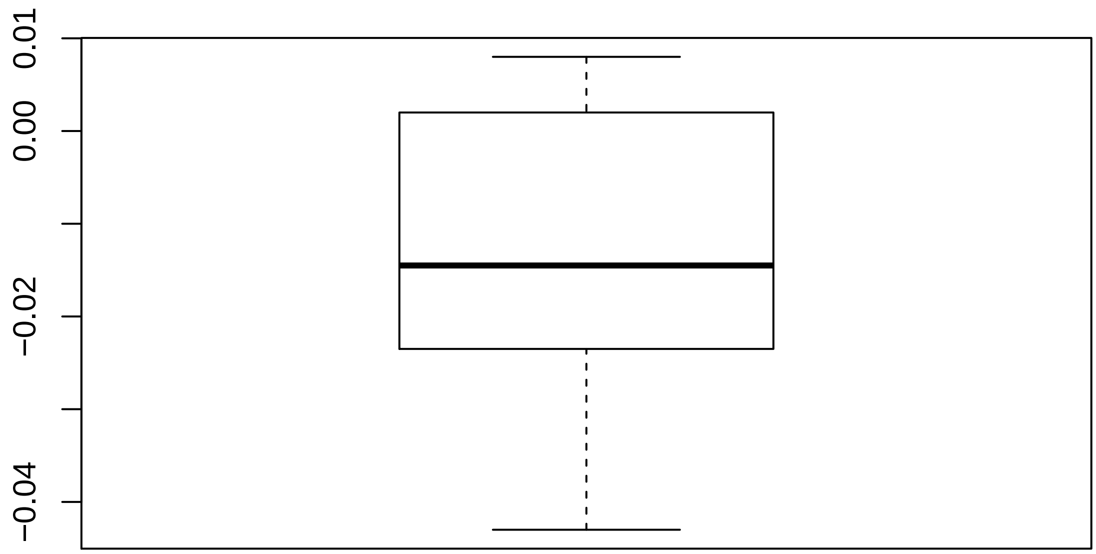
```
# Need to check assumptions for t-dist
n <- length(diff); n
```

```
## [1] 12
```

```
qqnorm(diff)
qqline(diff)
```




```r
boxplot(diff)
```




```r
shapiro.test(diff)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  diff
## W = 0.94325, p-value = 0.5413
```

p-value of shapiro wilk is .5413 so it is normally distributed

Compute test statistic

Null Hypothesis: Mu_d = 0 (difference equal to zero) Alternative Hyp: Mu_d < 0 (dominant hand less than nondominate)

```r
t <- ( mean(diff) - 0 ) / ( sd(diff) / sqrt(n) ); t
```

```
## [1] -2.775933
```

```r
# Method 1 - P-value
pvalue <- pt(t,n-1);
print("The p-value is"); pvalue
```

```
## [1] "The p-value is"

## [1] 0.00901731
```

```r
# Reject null

# Method 2 - Confidence Interval
upperbound <- mean(diff) + qt(0.95, n-1)* sd(diff)/sqrt(n); upperbound
```

```
## [1] -0.004648515
```

```r
print("The upperbound is: "); upperbound
```

```
## [1] "The upperbound is: "

## [1] -0.004648515
```

```r
# Method 3 - Use R with one sample test procedure
t.test(diff, mu=0, alternative = "less", conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  diff
## t = -2.7759, df = 11, p-value = 0.009017
## alternative hypothesis: true mean is less than 0
## 95 percent confidence interval:
##          -Inf -0.004648515
## sample estimates:
##   mean of x
## -0.01316667
```

```r
# Method 4 - Use R function with two sample test
t.test(hands$Dominant.Hand, hands$Nondominant.Hand, paired = T, alternative = "less", 0.95)
```

```
##
##  Paired t-test
##
## data:  hands$Dominant.Hand and hands$Nondominant.Hand
## t = -203.06, df = 11, p-value < 2.2e-16
## alternative hypothesis: true difference in means is less than 0.95
## 95 percent confidence interval:
##          -Inf -0.004648515
## sample estimates:
## mean of the differences
##             -0.01316667
```

Method 1 Decision

p-value is less than significance level so we reject null hypothesis

Method 2 Decision

Lower Bound: -infinity, Upperbound = -0.00464 Since it doesn't contain Zero, reject null

Changing tests to be two sided

```
t <- ( mean(diff) - 0 ) / ( sd(diff) / sqrt(n) ); t
```

```
## [1] -2.775933
```

```
# Method 1 - P-value
pvalue2 <- 2 * pt(t,n-1)
print("The p-value is"); pvalue2
```

```
## [1] "The p-value is"
```

```
## [1] 0.01803462
```

```
# Reject null

# Method 2 - Confidence Interval
lowerbound2 <- mean(diff) - qt(0.975, n-1)* sd(diff)/sqrt(n); lowerbound2
```

```
## [1] -0.02360627
```

```
upperbound2 <- mean(diff) + qt(0.975, n-1)* sd(diff)/sqrt(n); upperbound2
```

```
## [1] -0.002727063
```

```
# Method 3 - Use R with one sample test procedure
t.test(diff, mu=0, alternative = "two.sided", conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  diff
## t = -2.7759, df = 11, p-value = 0.01803
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.023606270 -0.002727063
## sample estimates:
##   mean of x
## -0.01316667
```

```
# Method 4 - Use R function with two sample test
t.test(hands$Dominant.Hand, hands$Nondominant.Hand, paired = T, alternative = "two.sided", 0.95)
```

```
##
##  Paired t-test
##
## data:  hands$Dominant.Hand and hands$Nondominant.Hand
## t = -203.06, df = 11, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0.95
## 95 percent confidence interval:
##  -0.023606270 -0.002727063
## sample estimates:
## mean of the differences
##             -0.01316667
```

Method 1 Decision

p-value is less than significance level so we reject null hypothesis

Method 2 Decision

Lower Bound: -0.0236, Upperbound = -0.0027 Since it doesn't contain Zero, reject null