Non-Parametric Hypothesis Testing

Ellie Bi & Terrie Kim

Introduction:

The purpose of this project is to perform a comprehensive analysis of data from non-normal distributions for one-sample, two-sample, and k-sample data, and use non-parametric tests to analyze them. The conclusions of the non-parametric tests are to be compared to the conclusions of parametric tests that are formed on the assumption of normality.

The primary focus is on hypothesis testing, comparing means, variances, distributions, and exploring the appropriateness of statistical methods for different datasets. Normal probability plots (Q-Q plots), histograms, and boxplots will be used to assess the underlying distribution of the data. Through this analysis, we aim to gain insights into the distribution of the data and evaluate the effectiveness of both parametric and non-parametric approaches in addressing specific hypotheses.

Sample 1 Analysis:

Sample 1 consisted of a singular sample meaning that we would be using a one-sample method. When first assessing the data, we plotted the data as a histogram to grab a better grasp and choose an appropriate testing method. The histogram was not normally distributed, and the Q-Q plot was heavy-tailed with then points at the end diverging from the Q-Q line. Some options for one-sample methods were the binomial test and creating a confidence interval. Confidence intervals only tell us if our population center is reasonable and would be more suited for two-sided evaluations.

We opted to proceed with a binomial test, a non-parametric test, to test for the median. We preferred to use the median as the population center because the histogram showed a skewed distribution. We chose to use a binomial test because binomial tests perform hypothesis testing based on the median. We conducted a hypothesis test where the null hypothesis is that $\theta_{0.5}=0$ and the alternative is that $\theta_{0.5}>0$. Since our sample size was 30, we decided that it was appropriate to use the normal approximation of the binomial distribution. We obtained a statistic of $B_{obs}=27$, where B_{obs} is the number of data points greater than 0. Using the formula for Z score, we obtained a Z of 4.38178. From there, we found that the p-value for our binomial test was

0.00000588567. This means that we reject the null hypothesis, indicating that the data is statistically significant enough to conclude that the median is greater than 0.

Using the same data, we performed a t-test to compare our results. If we assume normality of the data, the p-value of the t-test is 0.01972, meaning we still reach the same conclusion of rejecting the null hypothesis. The difference in p-values can be attributed to the normality assumption made by the parametric t-test.

Sample 2 Analysis:

For sample 2, we decided to use two-sample methods because the data is split into 2 groups. Again, we plotted both groups as histograms and checked the Q-Q plot. The Q-Q plots for both Group 1 and Group 2 have the points at the end diverging from the Q-Q line, indicating that we cannot assume normality. Both the histogram and Q-Q plots indicate the need for a non-parametric test to assess the data and perform the hypothesis test. For both the test for difference of means and difference of variances, some possible two-sample methods are permutation test, Wilcoxon Rank-Sum Test, and Kruskal-Wallis. Since Kruskal-Wallis, when performed with two groups, is the same as the Wilcoxon Rank-Sum Test, we chose to perform the Wilcoxon Rank-Sum Test for difference of means, and a permutation test for difference of variances.

First, we decided to use the Wilcoxon Rank-Sum Test with null hypothesis $\mu_1=\mu_2$ and alternative hypothesis $\mu_1<\mu_2$, just so we could compare and validate the p-value from the permutation test. First, we combined both groups of data and ranked them. We found ties in the data and found it appropriate to use expected value and variance to calculate the p-value. We tested with the sum of ranks associated with Group 1. Since the sample size is 30, the sample is large enough so that we can use the normal approximation to find the p-value. We carried out the calculations necessary for the Wilcoxon Rank-Sum Test which included finding μ, σ^2, expected W, variance W, W_{obs}, and found the p-value. Our W_{obs} is equal to the sum of the ranks from Group 1 which was 713. The p-value of the Wilcoxon Rank-Sum Test is 0.001409444. This means that we can reject the null hypothesis, indicating that the data is statistically significant enough to conclude that the mean of Group 2 is greater than the mean of Group 1.

We compared our results to a two-sample t-test which gave us a p-value of 0.002426. The t-test reaches the same conclusion as both the permutation test and the Wilcoxon Rank-Sum test of rejecting the null hypothesis. Although the p-value of the t-test allows for the conclusion to be

the same as the non-parametric tests, the p-value of the t-test is still greater than that of the non-parametric test. This is because the t-test assumes the population is approximately normal, while the two non-parametric tests do not make this assumption. The data has outliers and ties in ranks, both of which would typically not occur in a normally distributed population. The non-parametric tests are more likely to detect the true alternative, which is why the p-values for the non-parametric tests are smaller than the p-value of the t-test.

To test the difference of variances with null hypothesis $\sigma_1=\sigma_2$ and alternative hypothesis $\sigma_1\neq\sigma_2$, we conducted a permutation test. We visualized the data with boxplots and obtained the deviance for each group. To do this, we subtracted the median from each data point in the groups. We found an observed RMD of 0.3288561. This is calculated by dividing the means of the absolute values of the deviances. Then we simulated permutations and repeated the process for each permutation, collecting the RMDs of each permutation. To get the p-value, we plotted the permutations to choose a tail. Then we found the proportion of permuted RMDs that were less than or equal to our observed RMD and multiplied it by 2 to represent our two-sided test. The p-value of the Permutation Test on Deviance is 0.0008. Since the p-value is less than 0.05, we reject the null hypothesis, indicating that the data is statistically significant enough to suggest a difference in variances.

We compared our results to a parametric two-sample F-test. The p-value of the F-test is 0.0000000004319. Since it is less than 0.05, we reject the null hypothesis when we assume normality of the data. We are able to reach the same conclusions with both tests. The difference in p-values between the non-parametric and parametric test is again attributed to the assumption of normality that the parametric test makes.

Sample 3 Analysis:

For sample 3, which consisted of data from two distributions, we performed a Kolmogorov-Smirnov Test with a null hypothesis of $F_1(x)=F_2(x)$ and an alternative hypothesis of $F_1(x)\leq F_2(x)$. We plotted the data using histograms and Q-Q plot and determined that the data were not normally distributed, meaning we continue with Kolmogorov-Smirnov since it is a non-parametric test.

First, we combined the data from both distributions to start our analysis on finding the absolute differences. Then, for each data point from our combined data, we found the proportion of data from Distribution 1 and Distribution 2 that were less than the data point. We followed this

by finding the absolute difference between the proportions of that we found at each data point. Our observed K-S statistic was the max value from the absolute differences which was 0.4. To find the p-value, we simulated permutations and performed the test for each permutation, collecting the K-S statistics for each. Then, we plotted the permutation and used our observed K-S statistic to determine the tail needed to find the p-value. The p-value we obtained was the proportion of permuted K-S statistics greater than or equal to the observed K-S statistic. Since the p-value of the Kolmogorov-Smirnov test was 0.015, we rejected the null hypothesis, which means that the data is statistically significant enough to indicate a difference between distributions.

Sample 4 Analysis:

Sample 4 consisted of four different groups of data for which we wanted to perform a hypothesis test with null hypothesis $\mu_1=\mu_2=\mu_3=\mu_4$ against the alternative hypothesis that not all of the means were equal. Like with all previous data sets, we used the histograms and Q-Q plots to determine that the data were not normally distributed, resulting in the need for a non-parametric test. Possible k-sample methods are Kruskal-Wallis, permutation test, and multiple comparisons. Since we do not need to find out exactly which means are not equal, a multiple comparisons test is not necessary. We chose to perform the Kruskal-Wallis test which can handle ties in data more efficiently and is computationally less time-consuming when compared to the permutation test.

We ranked all of the data and obtained the mean rank for each data group. To calculate our KW statistic, we set the multiplier to be equal to 1 over the variance of all ranks. Then, we multiplied the multiplier to the sum of n times the quantity, mean ranks for each data group minus (N+1)/2, squared. The KW statistic calculated was 9.520661. The p-value we obtained from using the KW statistic with the chi squared distribution is 0.02311257. Since the p-value is less than 0.05, we reject the null hypothesis, which means that the data is statistically significant enough to indicate that not all of the means from the data groups are equal.

When compared to an F-test from ANOVA, we get that the p-value from the F-test is 0.4016169, which means that under the assumption of normality, we fail to reject the null hypothesis. The difference in conclusion between the non-parametric and parametric test is again attributed to the assumption of normality that the parametric test makes.

Conclusion:

Depending on the distribution of the data, parametric and non-parametric tests will result in different conclusions. The use of non-parametric tests when the datasets are not normally distributed is essential in detecting the true alternatives, and providing a more specific p-value that will determine whether or not we can reject the null hypothesis. Therefore, in cases where the assumptions of normality are questionable, the use of non-parametric tests are important in determining the most accurate and appropriate conclusion for the hypothesis test.

For one-sample tests, the binomial test is an appropriate non-parametric test to use. For two-sample tests, permutation tests, the Wilcoxon Rank-Sum test, the Kolmogorov-Smirnov test, etc. are appropriate non-parametric tests to use. For k-sample tests, the Kruskal-Wallis test is an appropriate non-parametric test to use.

The choice of statistical method we used varied depending on the specific characteristics of each dataset, emphasizing the importance of considering both parametric and non-parametric approaches in statistical analyses.

This project demonstrates the significance of using non-parametric tests when the data is not normally distributed, and that each method is most advantageous depending on how many samples there are.

Code Appendix

Ellie Bi & Terrie Kim

2023-11-27

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.3.2
```

```
library(tidyverse)
```

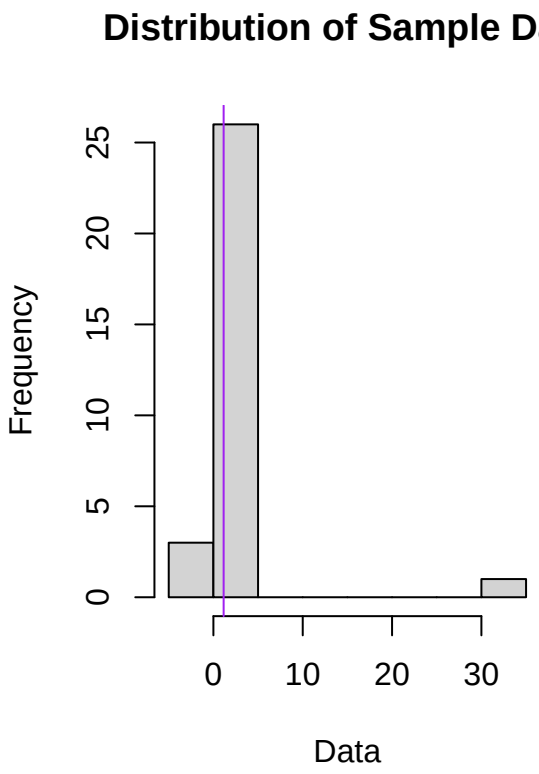
```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.3     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

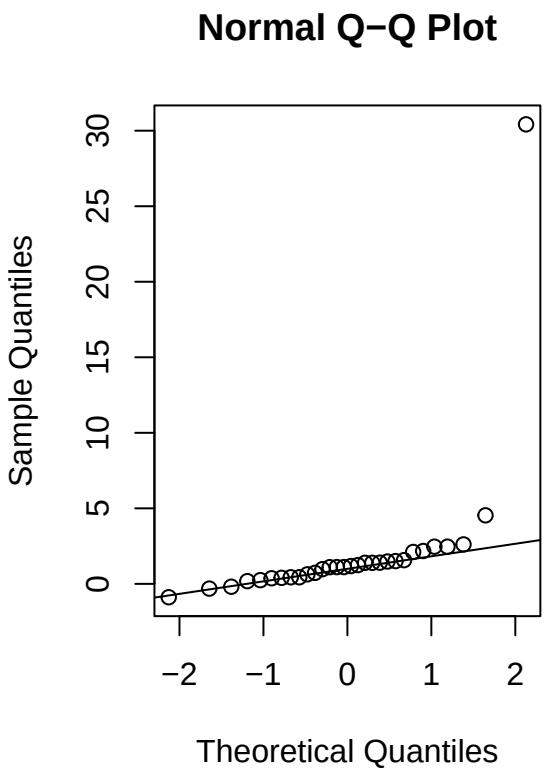
```
data <- read_excel('~/rstudio/sta104/term project data.xlsx', skip = 1)
```

P1 - One Sample Methods

```
sample1 <- data$sample
```

```
par(mfrow = c(1, 2))

hist(sample1, main = 'Distribution of Sample Data', xlab = 'Data')
abline(v = median(sample1), col = 'purple')

qqnorm(sample1)
qqline(sample1)
```





Binomial Test

```
#binomial test
n <- length(sample1)
B_obs <- sum(sample1 > 0); B_obs
```

```
## [1] 27
```

```
Z_b <- (B_obs - 0.5 * n) / sqrt(0.25 * n); Z_b
```

```
## [1] 4.38178
```

```
p_value <- pnorm(Z_b, lower.tail = FALSE); p_value
```

```
## [1] 5.88567e-06
```

T-Test

```
#t-test, still testing pop med
t_stat <- (mean(sample1) - 0) / (sd(sample1) / sqrt(length(sample1))); t_stat
```

```
## [1] 2.156908
```

```
p_value <- pt(abs(t_stat), df = length(sample1) - 1, lower.tail = FALSE); p_value
```

```
## [1] 0.01971874
```

```
t.test(sample1, alternative = c('greater'))
```

```
##
##  One Sample t-test
##
## data:  sample1
## t = 2.1569, df = 29, p-value = 0.01972
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  0.4549496       Inf
## sample estimates:
## mean of x
##  2.143567
```

P2 - Two-Sample Methods

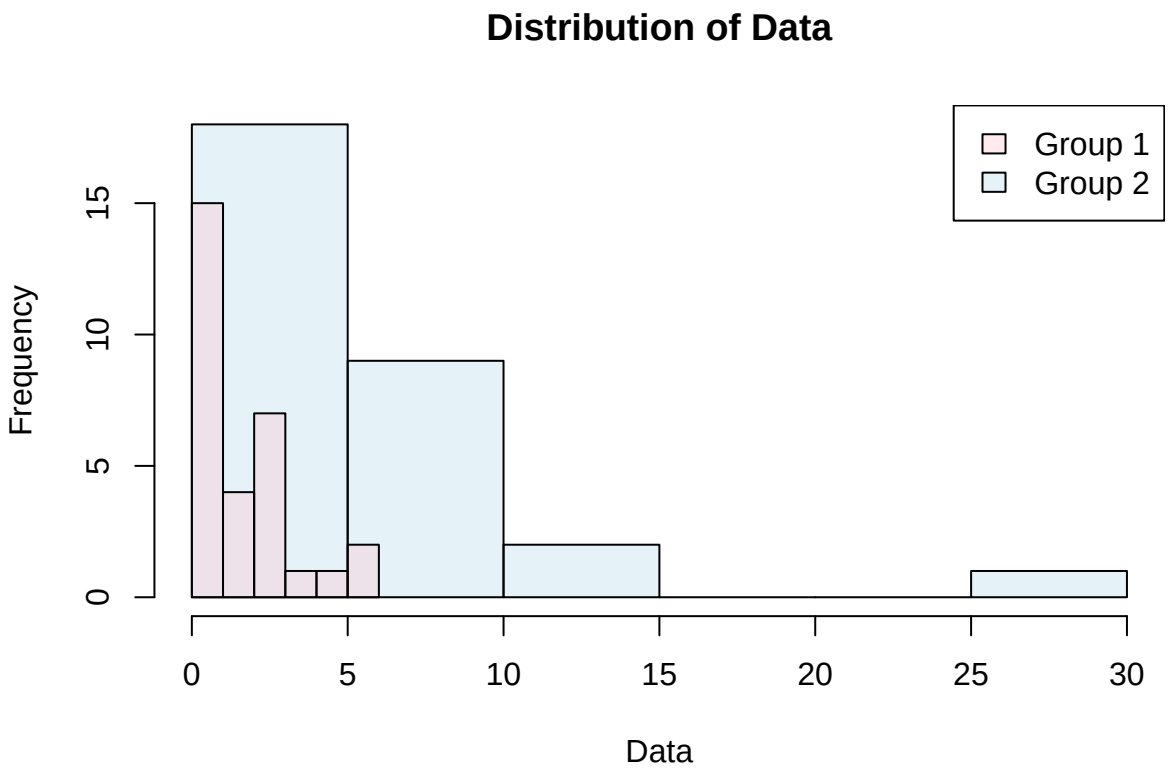
Means comparison

Permutation Test for Differences in Means

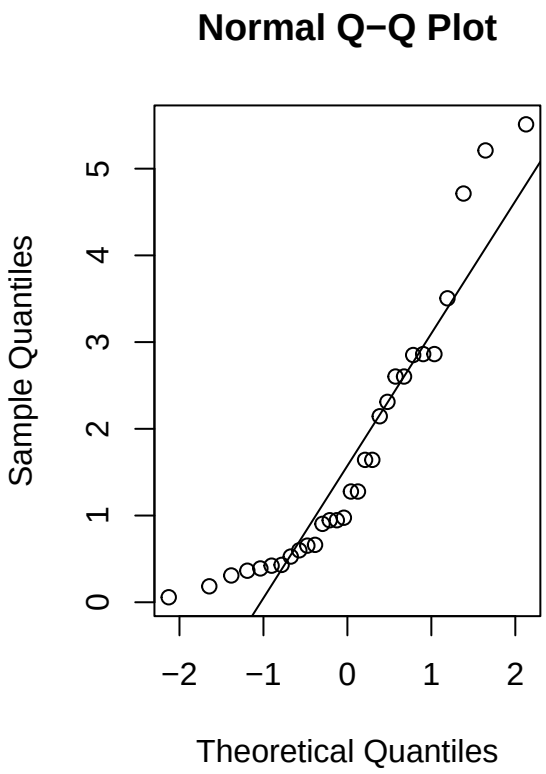
```
sample2 <- data.frame(Y = c(data$group1, data$group2),
                      group = c(rep(1,length(data$group1)),
                                rep(2,length(data$group2))))
```

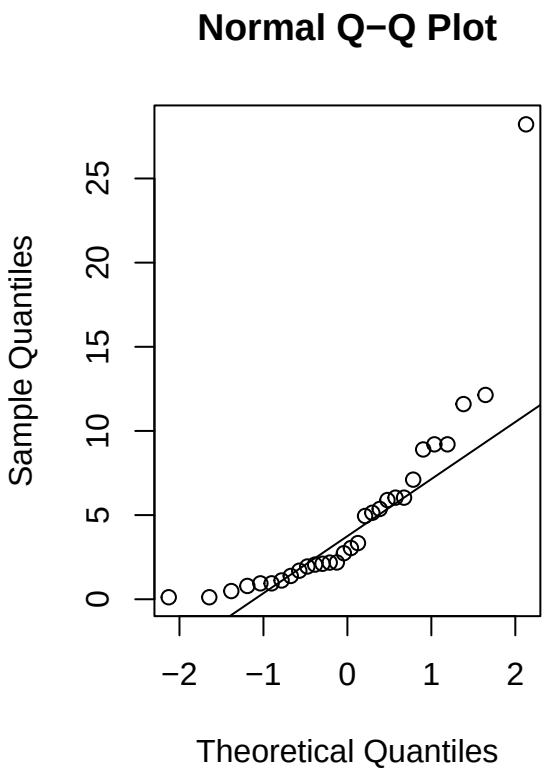
```
c1 <- rgb(173,216,230,max = 255, alpha = 80, names = "lt.blue")
c2 <- rgb(255,192,203, max = 255, alpha = 80, names = "lt.pink")
c3 <- rgb(220, 200, 255, max = 255, alpha = 80, names = "lt.purple")
c4 <- rgb(173, 216, 230, max = 255, alpha = 80, names = "lt.indigo")
```

```
hist(data$group2, col = c1, main = "Distribution of Data", xlab = "Data")
hist(data$group1, col = c2, add = TRUE)
legend("topright", legend = c("Group 1", "Group 2"), fill = c(c2, c1))
```

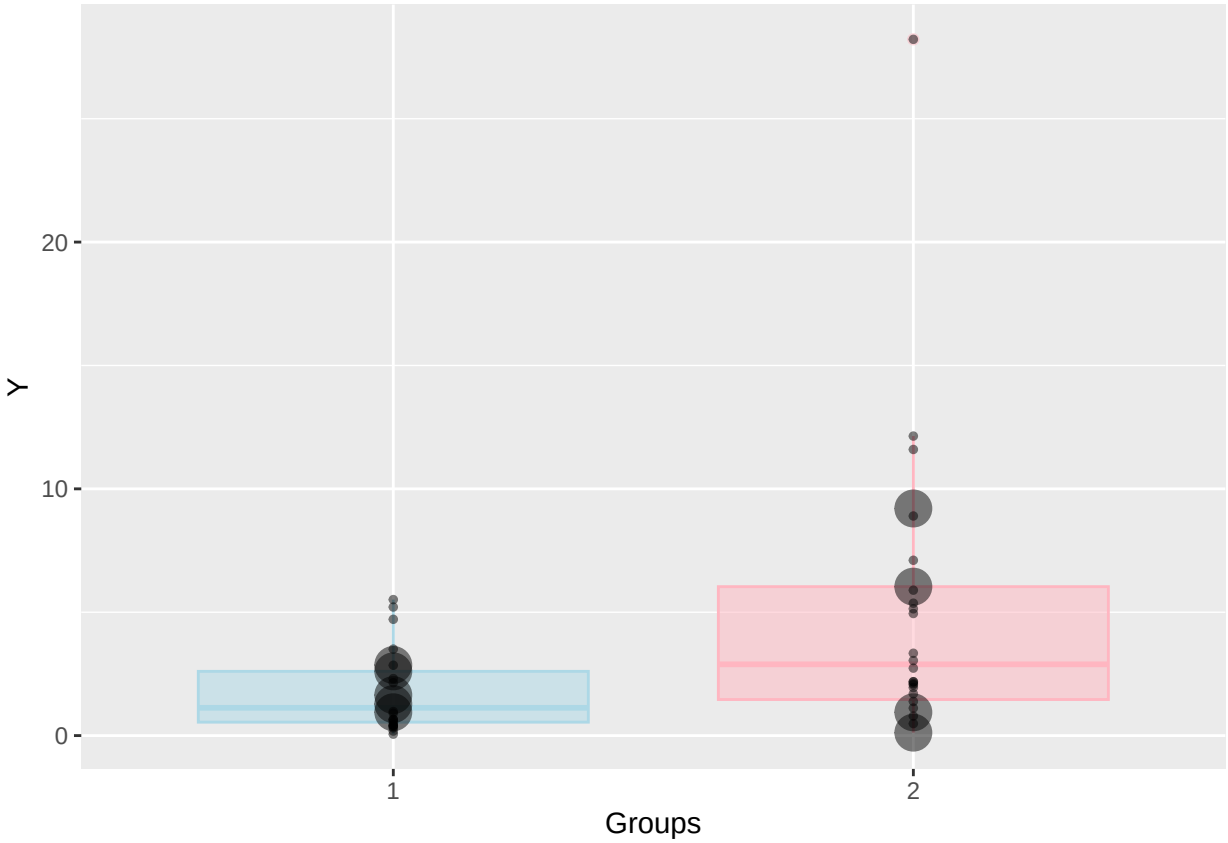



```
par(mfrow = c(1, 2))
qqnorm(data$group1)
qqline(data$group1)

qqnorm(data$group2)
qqline(data$group2)
```






```
ggplot(sample2, aes(x = factor(group), y = Y)) +
  geom_boxplot(aes(col = factor(group), fill = factor(group)), alpha = 0.5) +
  labs(x = "Groups") +
  scale_fill_manual(values = c("lightblue", "lightpink"), guide = FALSE) +
  scale_color_manual(values = c("lightblue", "lightpink")) +
  stat_sum(alpha = 0.5) +
  theme(legend.position = "none")
```



Wilcoxon Rank-Sum Test

```
ranks <- rank(c(data$group1, data$group2)); ranks
```

```
##  [1]  9.0 32.0 14.0 23.5 19.5 44.0  1.0  4.0 21.0 11.0  7.0 26.5 36.5  5.0  8.0
## [16] 50.0 12.0  6.0 40.5 16.0 13.0 45.0 39.0 35.0 48.0 26.5 36.5 40.5 23.5 19.5
## [31] 30.0 29.0 51.0 56.5 47.0 42.0 25.0 15.0 60.0 49.0 54.0 52.5 17.5 34.0 55.0
## [46] 58.0 33.0 10.0  2.5 22.0 31.0 38.0 43.0 28.0 46.0 52.5 17.5  2.5 56.5 59.0
```

```
g1_ranks <- ranks[1:30]
g2_ranks <- ranks[31:60]
```

```
mu <- sum(ranks) / length(ranks); mu
```

```
## [1] 30.5
```

```
sigma_sq <- (sum(ranks ** 2) / length(ranks)) - (mu ** 2); sigma_sq
```

```
## [1] 299.8417
```

```r
E_W <- length(data$group1) * mu; E_W
```

```
## [1] 915
```

```r
var_W <- (length(data$group1) * length(data$group2) * sigma_sq) /
  (length(ranks) - 1); var_W
```

```
## [1] 4573.856
```

```r
W_obs <- sum(g1_ranks); W_obs
```

```
## [1] 713
```

```r
p_value <- pnorm(-abs(W_obs - E_W) / sqrt(var_W), lower.tail = TRUE); p_value
```

```
## [1] 0.001409444
```

T-Test

```r
sp <- sqrt(((length(data$group1) - 1) * sd(data$group1)^2 +
             (length(data$group2) - 1) * sd(data$group2)^2) /
            (length(data$group1) + length(data$group2) - 2))

t_obs <- (mean(data$group1) - mean(data$group2)) /
  (sp * sqrt(1 / length(data$group1) + 1 / length(data$group2)))
t_obs
```

```
## [1] -3.018347
```

```r
pt(t_obs, df = length(data$group1) + length(data$group2) - 2, lower.tail = TRUE)
```

```
## [1] 0.001886204
```

```r
t.test(data$group1, data$group2, alternative = c('less'))
```

```
##
##  Welch Two Sample t-test
##
## data:  data$group1 and data$group2
## t = -3.0183, df = 33.249, p-value = 0.002426
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf -1.400068
## sample estimates:
## mean of x mean of y
##    1.7128    4.8989
```

Difference in Variance

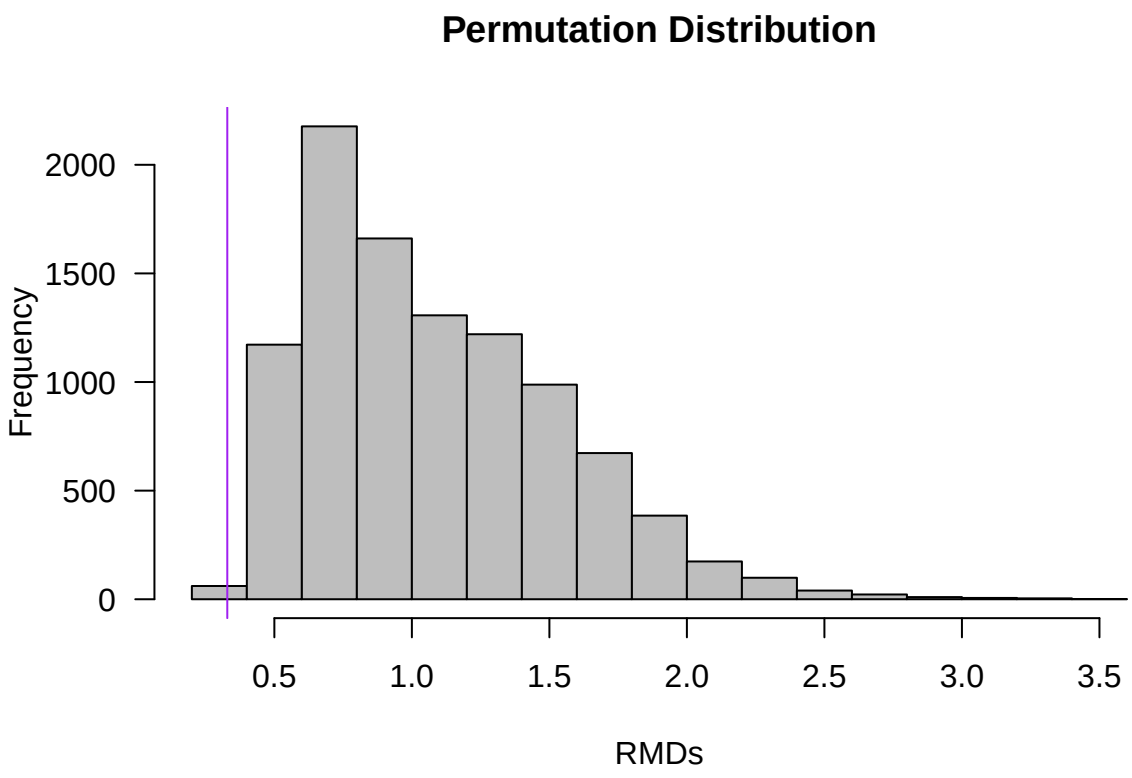
Permutation Test on Deviances

```
g1_dev <- data$group1 - median(data$group1)
g2_dev <- data$group2 - median(data$group2)

RMD_obs <- mean(abs(g1_dev)) / mean(abs(g2_dev)); RMD_obs
```

```
## [1] 0.3288561
```

```
set.seed(8)
permuted_RMDs <- numeric(10000)
for (i in 1:10000) {
  perm <- sample(nrow(sample2), replace = FALSE)
  p_data <- transform(sample2, group = group[perm])
  numer <- p_data$Y[p_data$group == 1] - median(p_data$Y[p_data$group == 1])
  denom <- p_data$Y[p_data$group == 2] - median(p_data$Y[p_data$group == 2])
  permuted_RMDs[i] <- mean(abs(numer)) / mean(abs(denom))
}
```

```
hist(permuted_RMDs, col = "gray", las = 1,
     main = "Permutation Distribution",
     xlab = "RMDs")
abline(v = RMD_obs, col = "purple")
```



```r
p_value <- 2 * mean(permuted_RMDs <= RMD_obs); p_value
```

```
## [1] 8e-04
```

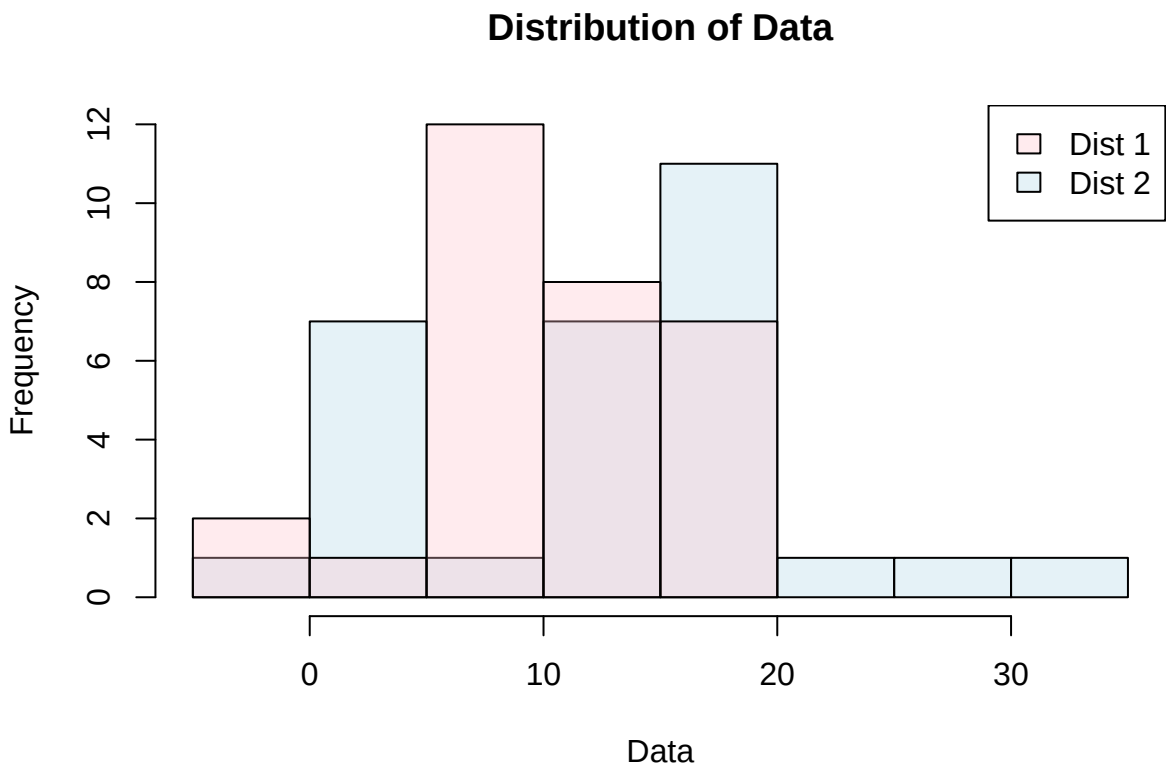
F-Test

```r
var.test(data$group1, data$group2, alternative = 'two.sided')
```

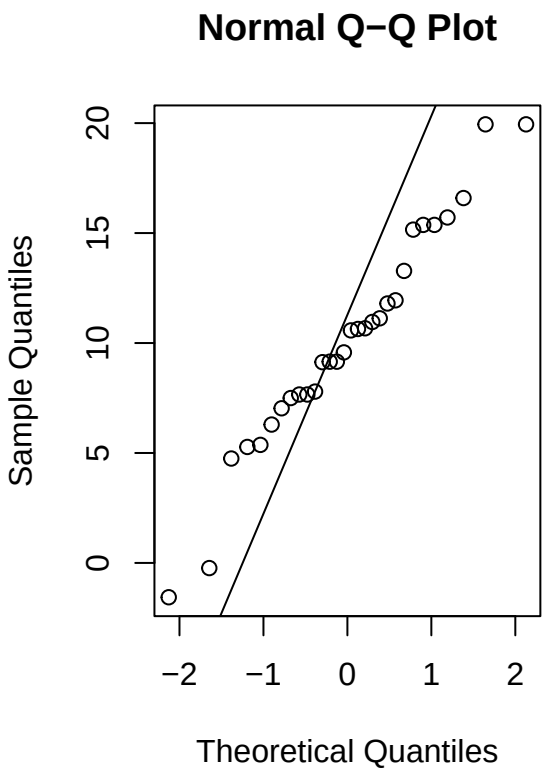
```
##
##  F test to compare two variances
##
## data:  data$group1 and data$group2
## F = 0.073651, num df = 29, denom df = 29, p-value = 4.319e-10
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.03505535 0.15474076
## sample estimates:
## ratio of variances
##         0.07365115
```

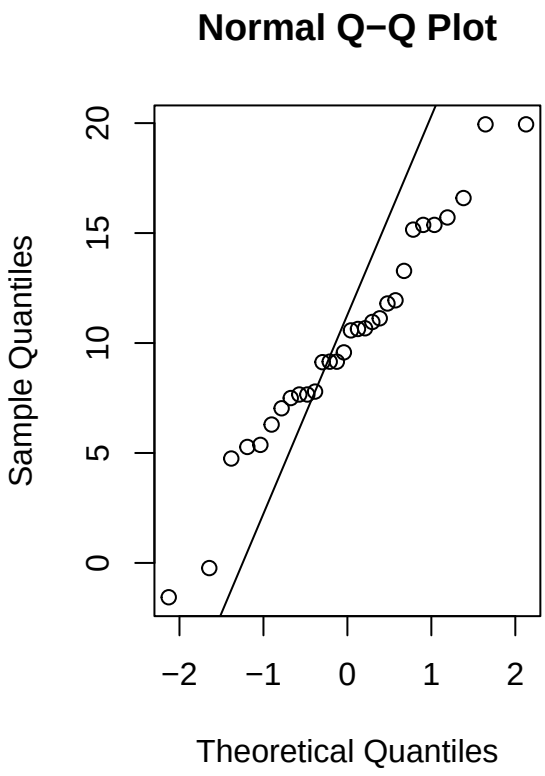
P3 - Two-Sample Method - Komolgorov-Smirnov Test

```r
hist(data$distrib2, col = c1, main = "Distribution of Data", ylim = c(0, 12),
     xlab = "Data")
hist(data$distrib1, col = c2, add = TRUE)
legend("topright", legend = c("Dist 1", "Dist 2"), fill = c(c2, c1))
```




```
par(mfrow = c(1, 2))
qqnorm(data$distrib1)
qqline(data$distrib2)

qqnorm(data$distrib1)
qqline(data$distrib2)
```






```
sample3 <- subset(data, select=c("distrib1", "distrib2"))
```

```
combined <- sort(c(sample3$distrib1, sample3$distrib2))

d1 <- c()
d2 <- c()

for (i in 1:60) {
  d1count = 0
  d2count = 0
  for (j in 1:30) {
    if (sample3$distrib1[j] < combined[i]) {
      d1count <- d1count + 1
    }
    if (sample3$distrib2[j] < combined[i]) {
      d2count <- d2count + 1
    }
  }
  d1[i] <- d1count / 30
  d2[i] <- d2count / 30
}

absdiff <- c()
for(i in 1:60){
  absdiff[i] <- abs(d1[i] - d2[i])
```

```
}
```

```
#visualizing the test
tab <- matrix(round(c(combined, d1, d2, absdiff), 3), ncol = 60, byrow = TRUE)

rownames(tab) <- c("data", "distribution 1", "distribution 2", "absolute difference")

as.table(tab)
```

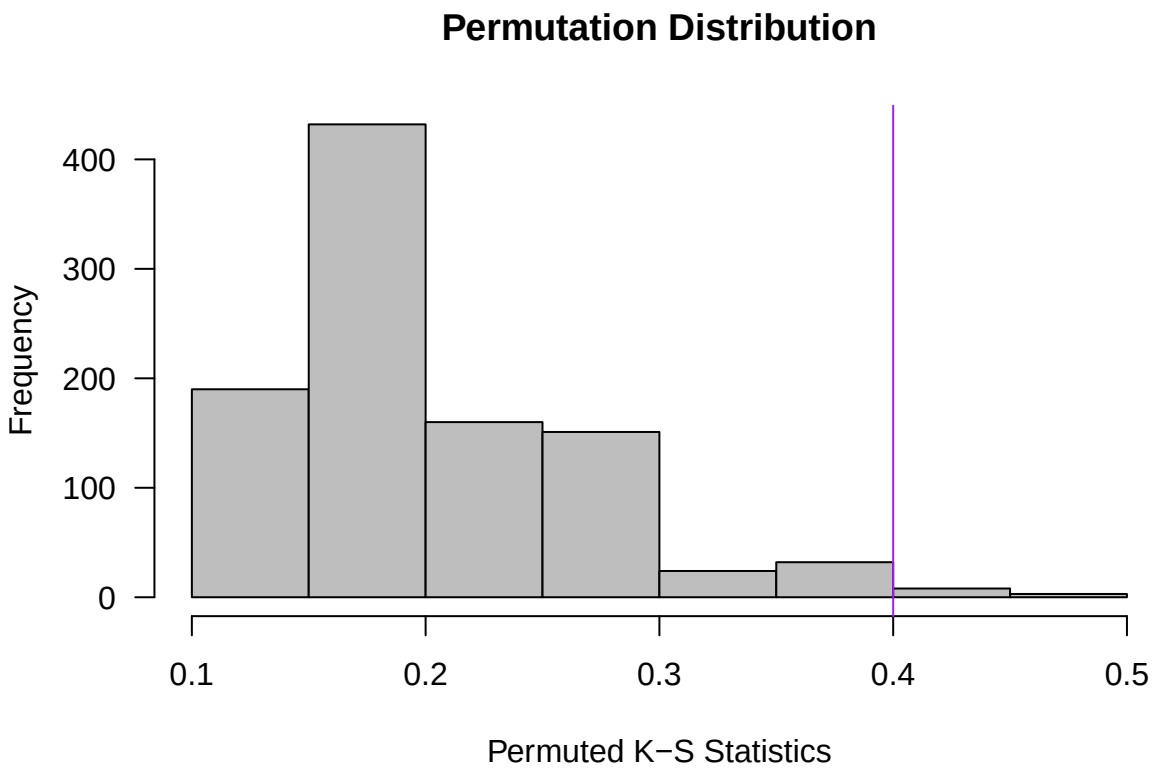
```
##                          A      B      C      D      E      F      G      H
## data                -3.345 -1.560 -0.234  0.556  1.116  1.866  2.112  3.788
## distribution 1       0.000  0.000  0.033  0.067  0.067  0.067  0.067  0.067
## distribution 2       0.000  0.033  0.033  0.033  0.067  0.100  0.133  0.167
## absolute difference  0.000  0.033  0.000  0.033  0.000  0.033  0.067  0.100
##                          I      J      K      L      M      N      O      P
## data                 4.748  4.926  4.926  5.274  5.368  5.948  6.292  7.032
## distribution 1       0.067  0.100  0.100  0.100  0.133  0.167  0.167  0.200
## distribution 2       0.200  0.200  0.200  0.267  0.267  0.267  0.300  0.300
## absolute difference  0.133  0.100  0.100  0.167  0.133  0.100  0.133  0.100
##                          Q      R      S      T      U      V      W      X
## data                 7.498  7.659  7.659  7.792  9.129  9.152  9.152  9.574
## distribution 1       0.233  0.267  0.267  0.333  0.367  0.400  0.400  0.467
## distribution 2       0.300  0.300  0.300  0.300  0.300  0.300  0.300  0.300
## absolute difference  0.067  0.033  0.033  0.033  0.067  0.100  0.100  0.167
##                          Y      Z     A1     B1     C1     D1     E1     F1
## data                10.430 10.574 10.639 10.670 10.957 11.124 11.798 11.942
## distribution 1       0.500  0.500  0.533  0.567  0.600  0.633  0.667  0.700
## distribution 2       0.300  0.333  0.333  0.333  0.333  0.333  0.333  0.333
## absolute difference  0.200  0.167  0.200  0.233  0.267  0.300  0.333  0.367
##                         G1     H1     I1     J1     K1     L1     M1     N1
## data                12.487 12.552 12.832 13.280 13.393 13.832 13.974 15.157
## distribution 1       0.733  0.733  0.733  0.733  0.767  0.767  0.767  0.767
## distribution 2       0.333  0.367  0.400  0.433  0.433  0.467  0.500  0.533
## absolute difference  0.400  0.367  0.333  0.300  0.333  0.300  0.267  0.233
##                         O1     P1     Q1     R1     S1     T1     U1     V1
## data                15.263 15.372 15.372 15.710 15.812 16.593 16.672 16.834
## distribution 1       0.800  0.800  0.800  0.867  0.900  0.900  0.933  0.933
## distribution 2       0.533  0.567  0.567  0.567  0.567  0.600  0.600  0.633
## absolute difference  0.267  0.233  0.233  0.300  0.333  0.300  0.333  0.300
##                         W1     X1     Y1     Z1     A2     B2     C2     D2
## data                16.918 16.918 17.539 17.539 17.643 18.815 18.815 19.944
## distribution 1       0.933  0.933  0.933  0.933  0.933  0.933  0.933  0.933
## distribution 2       0.667  0.667  0.733  0.733  0.800  0.833  0.833  0.900
## absolute difference  0.267  0.267  0.200  0.200  0.133  0.100  0.100  0.033
##                         E2     F2     G2     H2
## data                19.944 20.711 26.947 30.529
## distribution 1       0.933  1.000  1.000  1.000
## distribution 2       0.900  0.900  0.933  0.967
## absolute difference  0.033  0.100  0.067  0.033
```

```
ks_stat <- max(absdiff); ks_stat
```

```
## [1] 0.4
```

```r
set.seed(8)
p <- numeric(1000)
for (x in 1:1000) {
  permut <- sample(combined)
  pd1 <- permut[1:30]
  pd2 <- permut[31:60]

  d1 <- numeric(60)
  d2 <- numeric(60)

  for (i in 1:60) {
    d1count=0
    d2count=0
    for (j in 1:30) {
      if (pd1[j] < permut[i]) {
        d1count <- d1count + 1
      }
      if (pd2[j] < permut[i]) {
        d2count <- d2count + 1
      }
    }
    d1[i] <- d1count/30
    d2[i] <- d2count/30
  }

  absdiff <- numeric(60)
  for (i in 1:60) {
    absdiff[i] <- abs(d1[i] - d2[i])
  }

  ks_perm <- max(absdiff)
  p[x] <- ks_perm
}
```

```r
hist(p, col = "gray", las = 1,
     main = "Permutation Distribution",
     xlab = 'Permuted K-S Statistics')
abline(v = ks_stat, col = "purple")
```

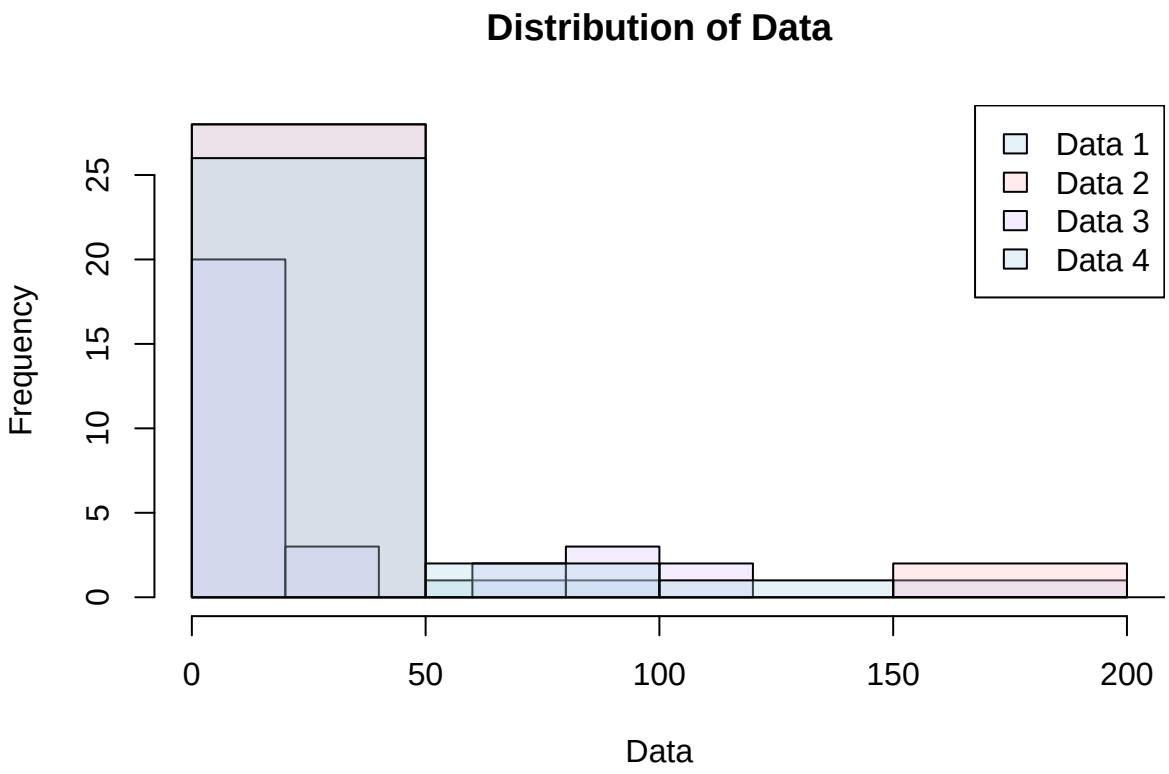



```
p_value <- mean(p >= ks_stat); p_value
```

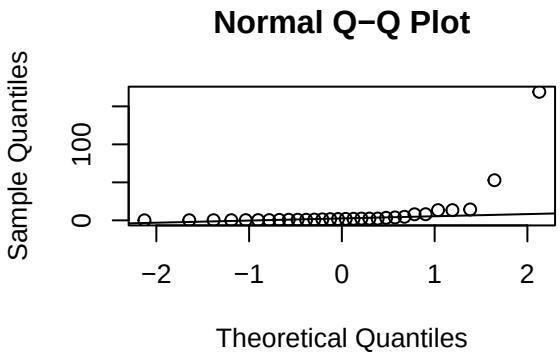
```
## [1] 0.015
```

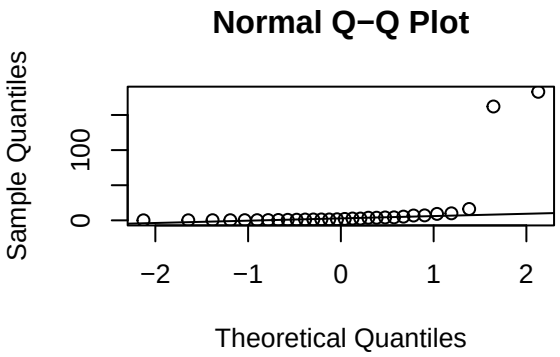
P4 - k-Sample Methods

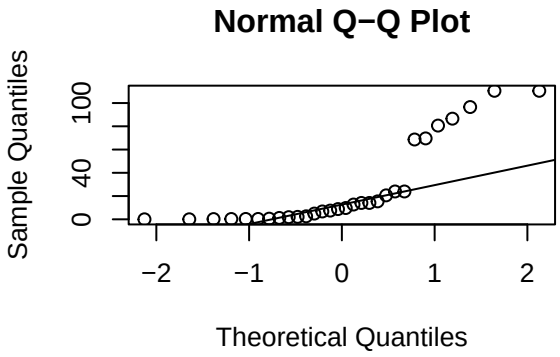
```
hist(data$cat1, col = c1, main = "Distribution of Data", xlab = "Data")
hist(data$cat2, col = c2, add = TRUE)
hist(data$cat3, col = c3, add = TRUE)
hist(data$cat4, col = c4, add = TRUE)
legend("topright", legend = c("Data 1", "Data 2", "Data 3", "Data 4"),
       fill = c(c1, c2, c3, c4))
```

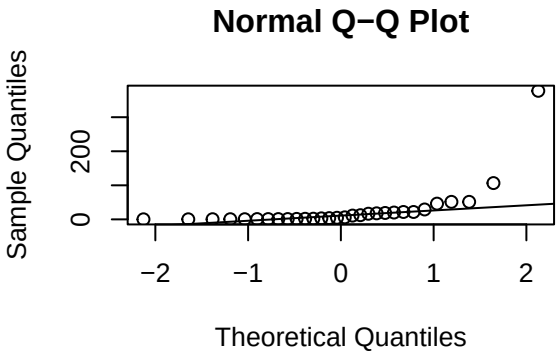



```
par(mfrow = c(2, 2))
qqnorm(data$cat1)
qqline(data$cat1)

qqnorm(data$cat2)
qqline(data$cat2)

qqnorm(data$cat3)
qqline(data$cat3)

qqnorm(data$cat4)
qqline(data$cat4)
```










```r
sample4 <- data.frame(Y = c(data$cat1, data$cat2, data$cat3, data$cat4),
                      group = c(rep('cat1',length(data$group1)),
                                rep('cat2',length(data$group2)),
                                rep('cat3',length(data$group2)),
                                rep('cat4',length(data$group2))))
head(sample4)
```

```
##         Y group
## 1   0.172  cat1
## 2   1.736  cat1
## 3 169.179  cat1
## 4   0.778  cat1
## 5   1.027  cat1
## 6   1.315  cat1
```

Kruskal-Wallis test

```r
ranks <- rank(sample4$Y)
cat1_R <- mean(ranks[1:30])
cat2_R <- mean(ranks[31:60])
cat3_R <- mean(ranks[61:90])
cat4_R <- mean(ranks[91:120])
mean(ranks)
```

```
## [1] 60.5
```

```
# since there are ties
multiplier <- 1 / var(ranks)
fraction <- (length(sample4$Y) + 1) / 2
KW_stat <- multiplier * ((30 * (cat1_R - fraction)^2) +
                         (30 * (cat2_R - fraction)^2) +
                         (30 * (cat3_R - fraction)^2) +
                         (30 * (cat4_R - fraction)^2))
KW_stat
```

```
## [1] 9.520661
```

```
1 - pchisq(KW_stat, 3)
```

```
## [1] 0.02311257
```

F-Test, ANOVA

```
anova <- summary(aov(Y ~ group, data = sample4)); anova
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## group         3   6735    2245   0.987  0.402
## Residuals   116 263899    2275
```

```
MST <- anova[[1]][1, 3]
MSE <- anova[[1]][2, 3]

f_stat <- MST / MSE
p_value <- 1 - pf(f_stat, 3, 116); p_value
```

```
## [1] 0.4016169
```