

This is an example of the Permutation Test and Bootstrapping to find confidence intervals and do a hypothesis test. We use the Megasonics data.. A QQ-Plot shows that the sample is not normal, so we look at subsets of the observations as a proxy for subsets of the underlying distribution. We wish to test the hypothesis that the megasonic cleaning results in fewer defects in a silicon wafer than the standard cleaning. From the t -test, even at the $\alpha = .10$ level, the data does not show that the mean defects of standard cleaning is significantly more than the mean defects of megasonic cleaning. In fact, the t -test one-sided 95% confidence interval is $-112.4 < \mu_S - \mu_M$. The difference is not significant at $\alpha = .05$. The two-sided 95% confidence interval would have been $-165.6 < \mu_S - \mu_M < 409.2$.

The permutation test is the most computationally intensive method for testing the hypotheses. In our case, there are $n_S = 10$ and $n_M = 5$ wafers cleaned in the standard and megasonic ways. All together, there were $n_S + n_M$ wafers cleaned by one method or the other. Under the null hypothesis, there is no difference in the number of defects on the wafer for either cleaning method. The only variation in the mean of a sample of n_M wafers chosen for megasonic cleaning from the $n_M + n_S$ total wafers comes from which on the n_M wafers were chosen. All together there are $\binom{n_M+n_S}{n_M}$ choices, which in our case is $\binom{15}{5} = 3003$. Thus for all combinations of the $n_M + n_S$ wafers taken n_M at a time, we take the difference of the mean of those n_M defect and the mean of the n_S wafers. Note that this is a linear function of the sum of the megasonic defects. Let $\{x_1, \dots, x_{n_S+n_M}\}$ denote the number of observed defects. Let $\pi = \{j_1, \dots, j_{n_M}\}$ be a subset of $T = \{1, 2, 3, \dots, n_M + n_S\}$. Let $S = M - \pi = \{k_1, \dots, k_{n_S}\}$ be the complementary subset.

$$D_\pi = \frac{1}{n_S} \sum_{i=1}^{n_S} x_{k_i} - \frac{1}{n_M} \sum_{i=1}^{n_M} x_{j_i} = \frac{1}{n_S} \sum_{i=1}^{n_S+n_M} x_i - \left(\frac{1}{n_S} + \frac{1}{n_M} \right) \left(\sum_{i=1}^{n_M} x_{j_i} \right) = a + b \sum_{i=1}^{n_M} x_{j_i}.$$

Consider the totality of all possible differences $\{D_\pi\}$ where $\pi \subset T$ runs over n_M member subsets. The permutation confidence interval is the α -quantile of the D_π 's. In the case of the cleaning data, the one-sided $\alpha = .05$ permutation confidence interval is $-267.6 < \mu_S - \mu_M$. The one-sided $\alpha = .10$ permutation confidence interval is $-165.9 < \mu_S - \mu_M$, so the difference is not significantly positive at this confidence level from the permutation test either. The two-sided 95% confidence interval would have been $-295.8 < \mu_S - \mu_M < 245.1$. The p -value for the one sided test is the proportion of all differences such that exceed the observed difference $D_\pi \geq 121.8$. This p -value works out to be 0.163.

The number of combinations becomes prohibitively large as n_M and n_S get large. Instead of looking at all combinations, we approximate with a bootstrap selection of B combinations, $\{\pi_1, \pi_2, \dots, \pi_B\}$ where repetition of combinations is allowed. In our example $B = 500$ for illustration purposes, but $B = 10,000$ is better when $\binom{n_M+n_S}{n_M}$ is humongous. The bootstrap confidence intervals are the quantiles of the $\{D_{\pi_1}, \dots, D_{\pi_B}\}$. Our bootstrap one-sided $\alpha = .05$ confidence interval turned out to be $-237.3 < \mu_S - \mu_M$. The one-sided $\alpha = .10$ bootstrap confidence interval is $-158.1 < \mu_S - \mu_M$, so the difference is not significantly positive at this confidence level from the bootstrap test either. The two-sided 95% confidence interval would have been $-279.9 < \mu_S - \mu_M < 245.7$. The p value for the one-sided test is the proportion of bootstrapped differences such that exceed the observed difference $D_\pi \geq 121.8$. This p -value works out to be .202.

Data File Used in this Analysis:

```
# Math 3070 - 1      Megasonic Data          June 30, 2011
# Treibergs
#
# Data from Levine, Ramsey Smidt, "Applied Statistics for Engineers and
# Scientists," Prentice hall 2001.
# A study by C.W.Draper, "A Portable Nitrogen Purged Microenvironment"
# in Proceedings of Institute of Environmental Sciences, 1995,
# compared contamination of silicon wafer surfaces with particles and
# surface chemical oxidation. Ten wafers receiving a standard clean in a
# cleanroom were compared with five wafers receiving a megasonics clean
# and transported through a non-clean environment in the microenvironment.
# Defects were measured on 2 vernier patterns/wafer.
#
Defects Cleaning
53 Standard
193 Standard
113 Standard
640 Standard
800 Standard
140 Standard
85 Standard
658 Standard
140 Standard
140 Standard
26 Megsonic
90 Megsonic
546 Megsonic
90 Megsonic
120 Megsonic
```

R Session:

```
R version 2.11.1 (2010-05-31)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```

```
    Natural language support but running in an English locale
```

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```

```
Type 'demo()' for some demos, 'help()' for on-line help, or
```

'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

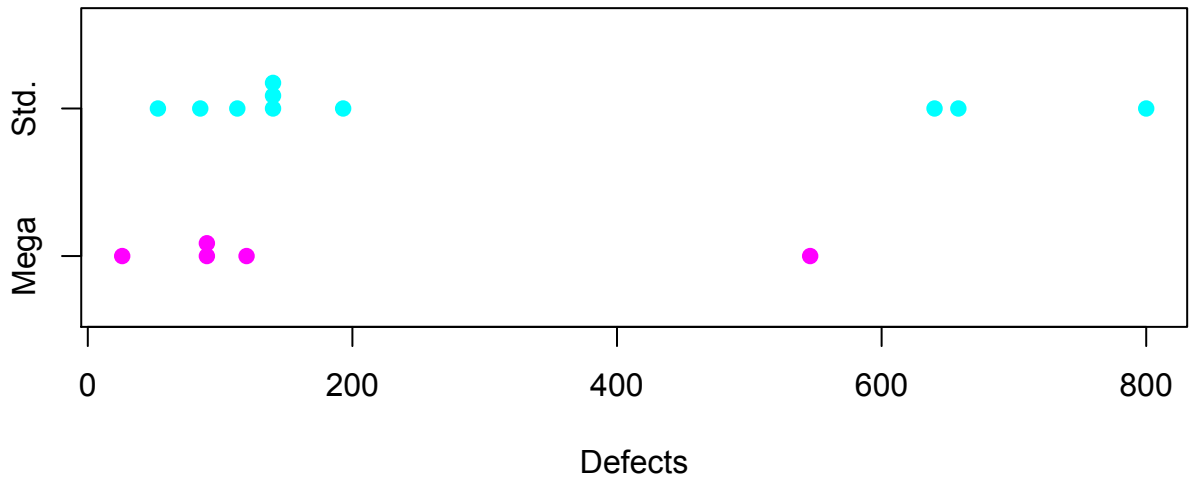
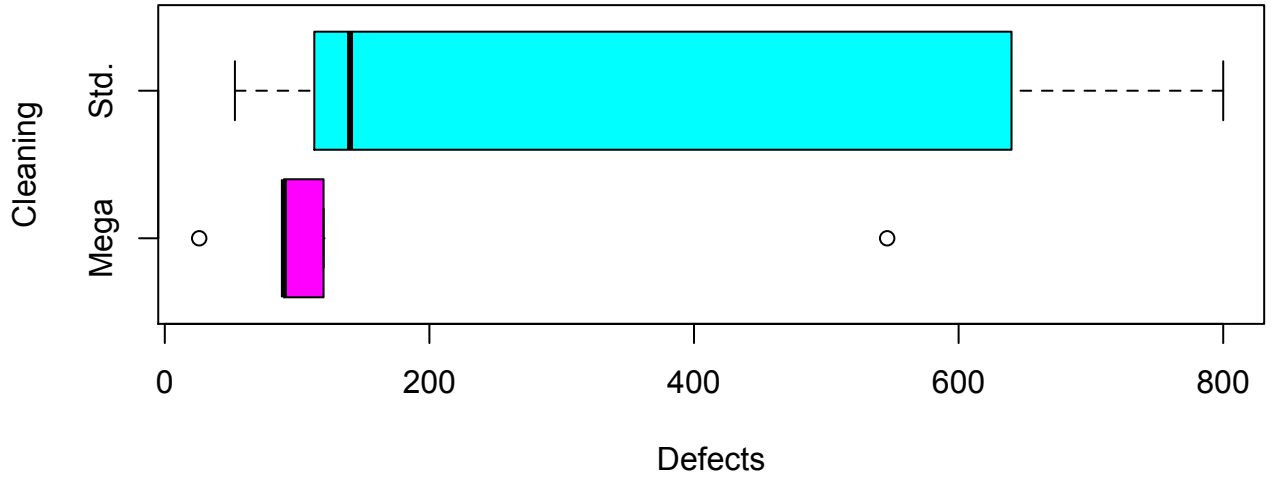
[R.app GUI 1.34 (5589) i386-apple-darwin9.8.0]

[Workspace restored from /home/1004/ma/treibergs/.RData]

```
> ##### ENTER DATA #####
> tt <- read.table("M3074MegasonicData.txt",header=TRUE)
> attach(tt)
> tt
  Defects Cleaning
1      53 Standard
2     193 Standard
3     113 Standard
4     640 Standard
5     800 Standard
6     140 Standard
7      85 Standard
8     658 Standard
9     140 Standard
10    140 Standard
11     26 Megasonic
12     90 Megasonic
13    546 Megasonic
14     90 Megasonic
15    120 Megasonic

> # Shorten the factor level names
> levels(Cleaning) <- c("Mega","Std.")
> # Make comparative boxplots and stripcharts
> layout(matrix(1:2,ncol=1))
> plot(Defects ~ Cleaning, horizontal = T, col = c("magenta","cyan"),
+ main="Megasonic vs. Standard Cleaning")
> stripchart(Defects ~ Cleaning, col = c("magenta", "cyan"), method = "stack",
+ ylim = c(.6,2.6), pch = 19)
> # M3074Megasonic1.pdf
>
> # The Megasonic Cleaning has fewer defects, but is it significantly less?
```

Megasonic vs. Standard Cleaning



```

> ### T-TEST FOR DIFFERENCE IN MEGASONIC CLEANING AND STANDARD CLEANING ###
> MC <- Defects[Cleaning=="Mega"]
> SC <- Defects[Cleaning=="Std."]
> t.test(SC,MC)

Welch Two Sample t-test

data: SC and MC
t = 0.9366, df = 10.654, p-value = 0.3697
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -165.5587  409.1587
sample estimates:
mean of x mean of y
   296.2    174.4

> t.test(SC,MC,alternative="greater")

Welch Two Sample t-test

data: SC and MC
t = 0.9366, df = 10.654, p-value = 0.1849
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -112.4366      Inf
sample estimates:
mean of x mean of y
   296.2    174.4

> ##### DIFFERENCE OF MEANS AS FUNCTION OF MEGASONIC SUM #####
> nS <- length(SC); nM <- length(MC)
> # number of combinations
> choose(nS+nM,nM)
[1] 3003

> sumD <- sum(Defects); sumD
[1] 3834
> # note that sum(SC)+ sum(MC) = sumD so
> mean(SC)-mean(MC)
[1] 121.8
> a <- sumD/nS; b <- -1/nS - 1/nM
> a + b*sum(MC)
[1] 121.8

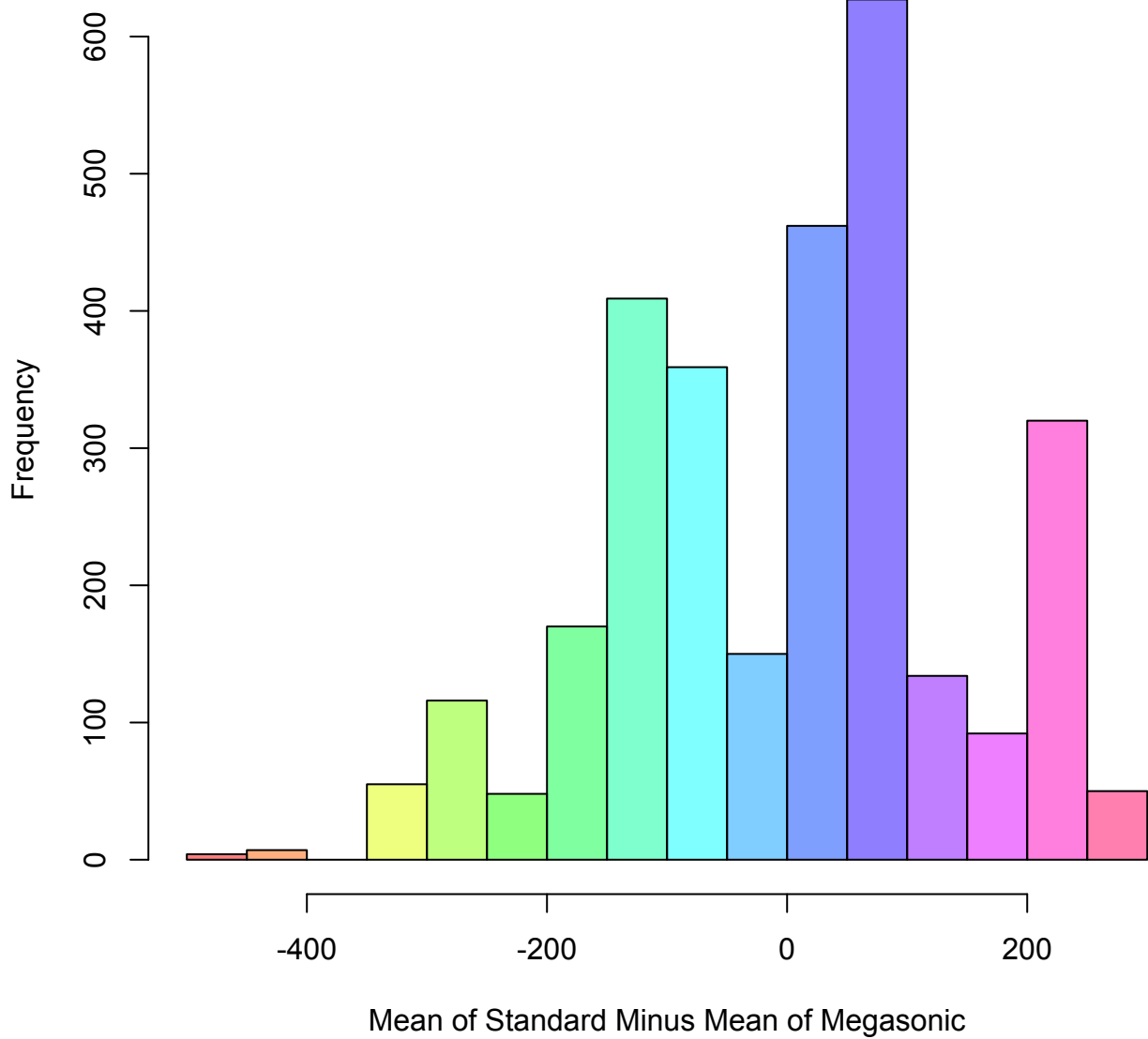
```

```

> ##### PLOT DISTRIBUTION OF ALL DIFFERENCES. FIRST HISTOGRAM #####
> dif <- function(x) a+b*sum(x)
> # combn(x,k,FUN=f) makes all combos of list x taken k
> # at a time and applies f to each combo.
> layout(1)
> hist(combn(Defects, 5, FUN = dif), col = rainbow(16,alpha=.5),
+ main = "Histogram of Differences from All Permutations",
+ xlab = "Mean of Standard Minus Mean of Megasonic")

```

Histogram of Differences from All Permutations



```

> ##### QUANTILES OF DIFFERENCES OF PERMUTATIONS #####
> difV <- combn(Defects,5,FUN=dif)
> quantile(difV)
  0%   25%   50%   75%  100%
-467.7 -107.4  25.2  83.7  280.2
>
> quantile(difV,c(.001,.0025,.005,.01,.025,.05,.1))
  0.1%  0.25%  0.5%   1%   2.5%   5%   10%
-451.7880 -436.0425 -337.7310 -317.7000 -295.8000 -267.6000 -165.9000
> quantile(difV,c(.9,.95,.975,.99,.995,.9975,.999))
  90%   95%  97.5%   99%  99.5%  99.75%  99.9%
212.7000 231.3000 245.1000 256.2000 263.7000 265.2000 271.7988

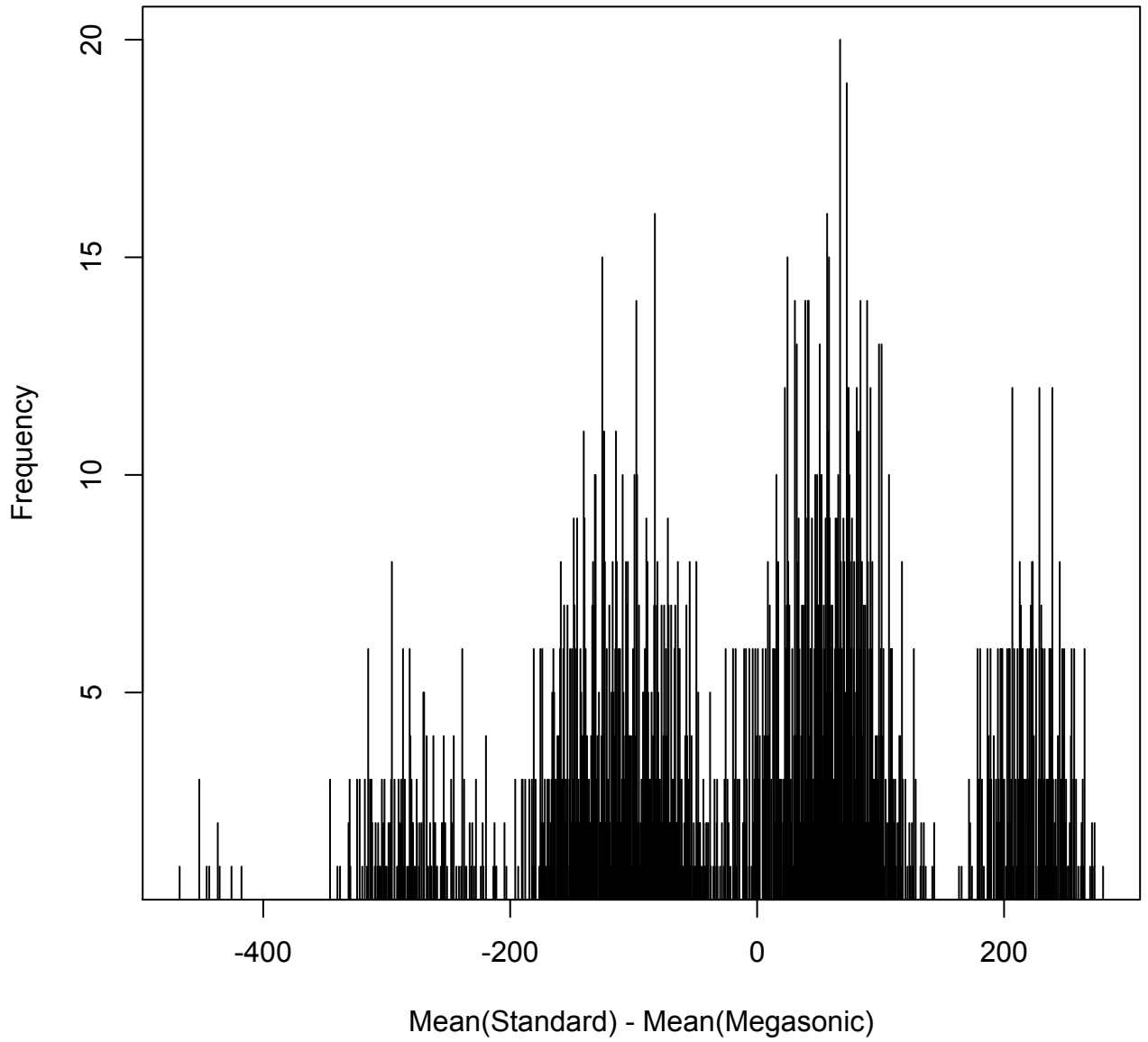
> ##### P-VALUE OF PERMUTATION TEST #####
> difT <- table(difV)
> length(difT)
[1] 776

> k <- 1
> while (as.numeric(names(difT[k])) < 121.8) k <- k+1
> difT[k-1]; difT[k]
120
  3
121.8
  1
> lp <- sum(difT[k:776])
> pvalue <- lp/3003; pvalue
[1] 0.1628372
>
>
> ##### TABULATE FREQUENCIES OF DIFFERENCES THAT OCCUR #####
>
> difT <- table(difV)
> length(difT)
[1] 776
> difT[1:10]
difV
-467.7 -451.8 -445.8 -443.7 -436.8 -435.3 -425.7 -417.6 -345.9 -339.9
  1     3     1     1     2     1     1     1     3     1
> difT[767:776]
difV
262.2 262.8 263.7 264.3 265.2 269.7 271.2 271.8 273.3 280.2
  1     2     3     1     6     1     2     1     2     1

> ##### PLOT FREQUENCIES OF DIFFERENCES THAT OCCUR #####
>
> plot(as.numeric(names(difT)), difT, type = "h",
+ main = "Frequencies of Differences Occuring from All Permutations",
+ xlab = "Mean(Standard) - Mean(Megasonic)", ylab = "Frequency")
> # M3074Megasonic3.pdf

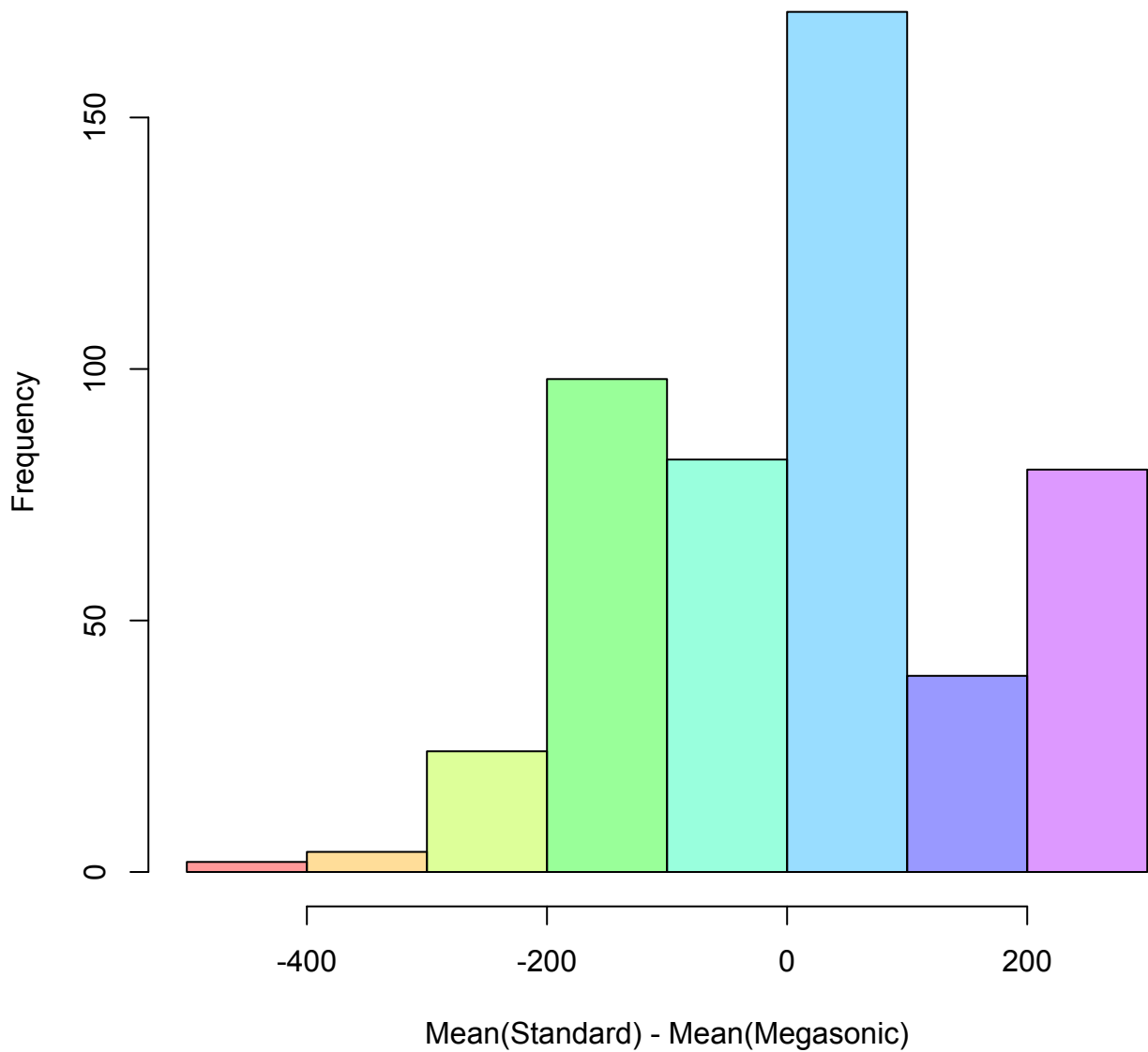
```

Frequencies of Differences Occuring from All Permutations




```
> ##### PLOT DISTRIBUTION OF BOOTSTRAPPED DIFFERENCES. FIRST HISTOGRAM #####
> B <- 500
> # Bootstrap sample. Choose B random combinations.
> difV2 <- replicate(B, a+b*sum(sample(Defects, 5, replace=FALSE)))
> hist(difV2, col = rainbow(9, alpha=.4),
+ main = "Histogram of Differences from Bootstrap Sample of Size 500",
+ xlab = "Mean(Standard) - Mean(Megasonic)")
> # M3074Megasonic4.pdf
```

Histogram of Differences from Bootstrap Sample of Size 500



```

> ##### QUANTILES OF DIFFERENCES OF PERMUTATIONS #####
>
> # Estimate CI using Bootstrap
>
> quantile(difV2,c(.001,.0025,.005,.01,.025,.05,.1))
      0.1%    0.25%    0.5%    1%    2.5%    5%    10%
-444.7521 -415.5593 -329.7030 -315.0270 -279.9000 -237.3000 -158.1000

> quantile(difV2,c(.9,.95,.975,.99,.995,.9975,.999))
      90%    95%    97.5%    99%    99.5%    99.75%    99.9%
220.8300 237.6300 245.7000 256.8540 264.7545 269.7150 275.7090
>
>
> ##### P-VALUE OF BOOTSTRAP TEST #####
>
> difT2 <- table(difV2)
> difT2[1:10]
difV2
-445.8 -443.7  -330 -329.4 -317.7  -315 -295.8 -286.8 -285.3 -281.4
      1      1      1      1      1      1      2      1      1      1
> > length(difT2)
[1] 319
> difT2[310:319]
difV2
247.2 250.2 254.7 256.2 256.8 262.2 264.3 265.2 271.2 280.2
      1      1      1      1      2      1      1      1      1      1
> sum(difT2)
[1] 500
> c(difT2[k-1], difT2[k])
117.9 121.8
      1      1
> sum(difT2)
[1] 500
> k <- 1
> while (as.numeric(names(difT2[k])) < 121.8) k <- k+1
> c(difT2[k-1], difT2[k])
117.9 121.8
      1      1
> sum(difT2[k:319])
[1] 101
> pvalue2 <- 101/500; pvalue2
[1] 0.202
>
> ##### PLOT FREQUENCIES OF DIFFERENCES THAT OCCUR IN BOOTSTRAP #####
>
> plot(as.numeric(names(difT2)), difT2, type = "h",
+ main = "Frequencies of Differences Occuring from Bootstrapping, B=500",
+ xlab = "Mean(Standard) - Mean(Megasonic)", ylab = "Frequency")
> # M3074Megasonic5.pdf

```

Frequencies of Differences Occuring from Bootstrapping, B=500

