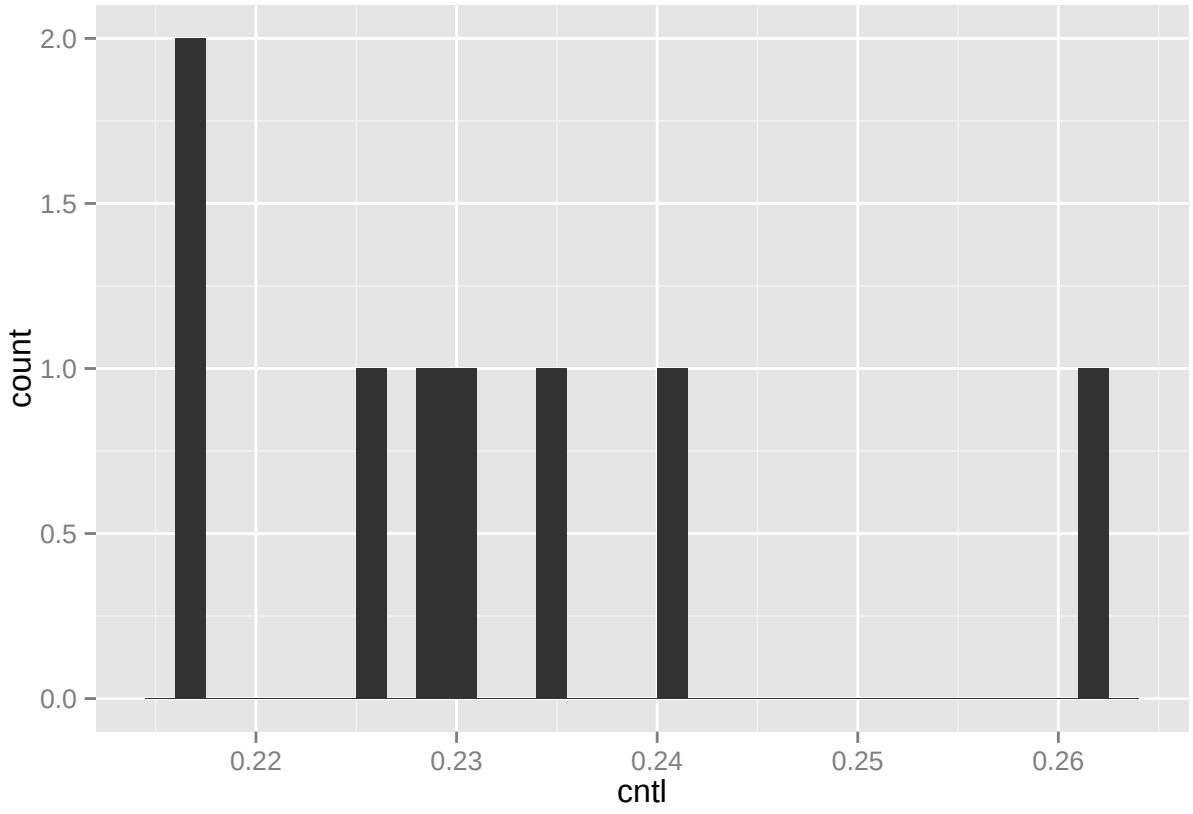
hw5.R

jiayuan

Mon Oct 26 12:51:44 2015

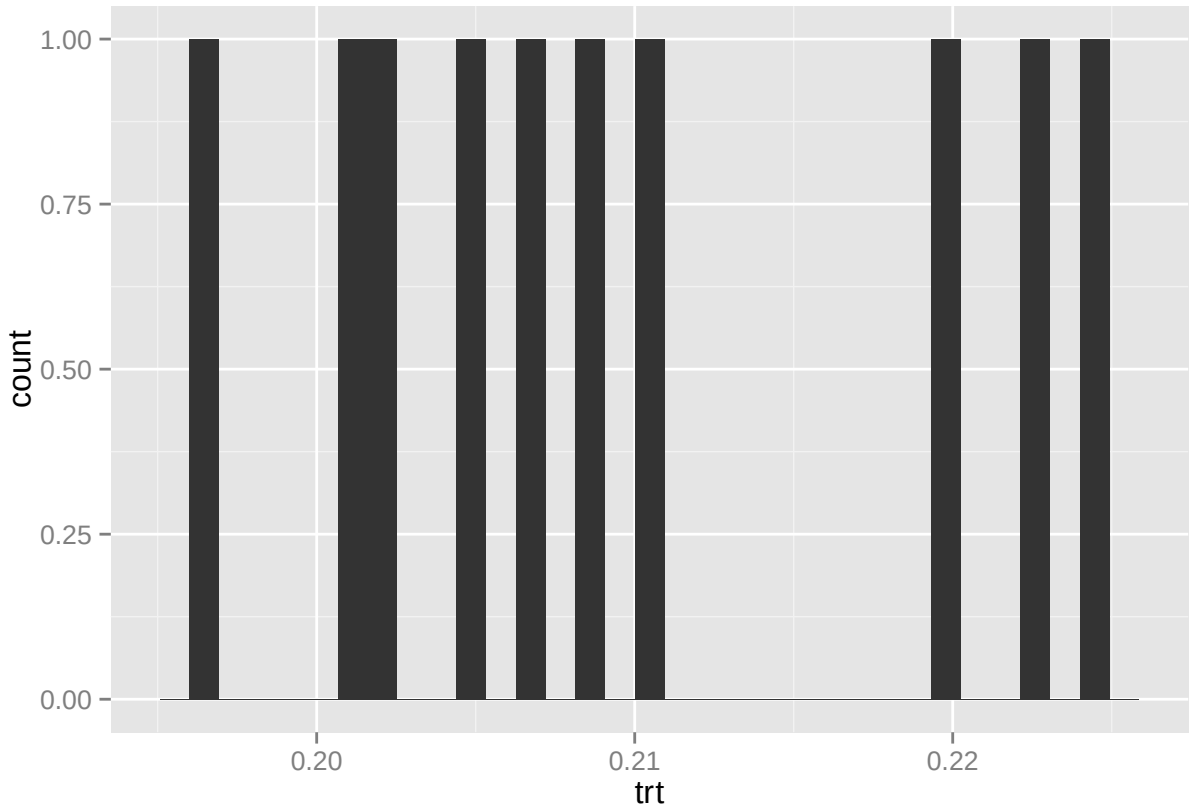
```
#4
library(ggplot2)
cntl <- c(0.225, 0.262, 0.217, 0.240,0.230, 0.229, 0.235, 0.217)
trt <-c(0.209, 0.205, 0.196, 0.210, 0.202,0.207, 0.224, 0.223, 0.220, 0.201)

qplot(cntl)
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```

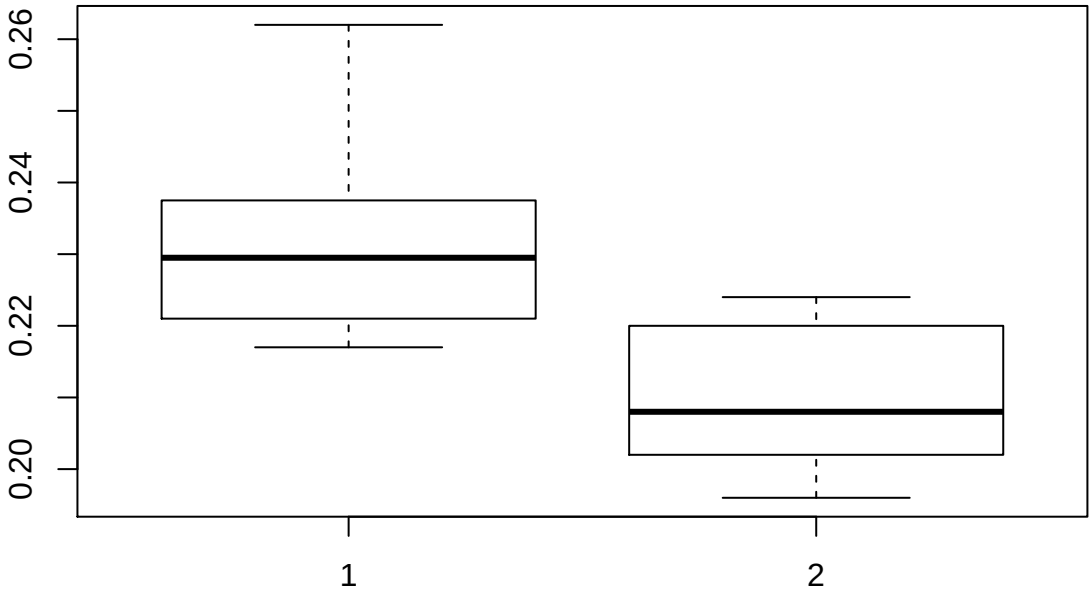



```
qplot(trt)
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```




```
boxplot(cntl,trt)
```




```
summary(cntl)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2170  0.2230  0.2295  0.2319  0.2362  0.2620
```

```r
summary(trt)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1960  0.2028  0.2080  0.2097  0.2175  0.2240
```

```r
# two sample t test
t.test(cntl,trt)
```

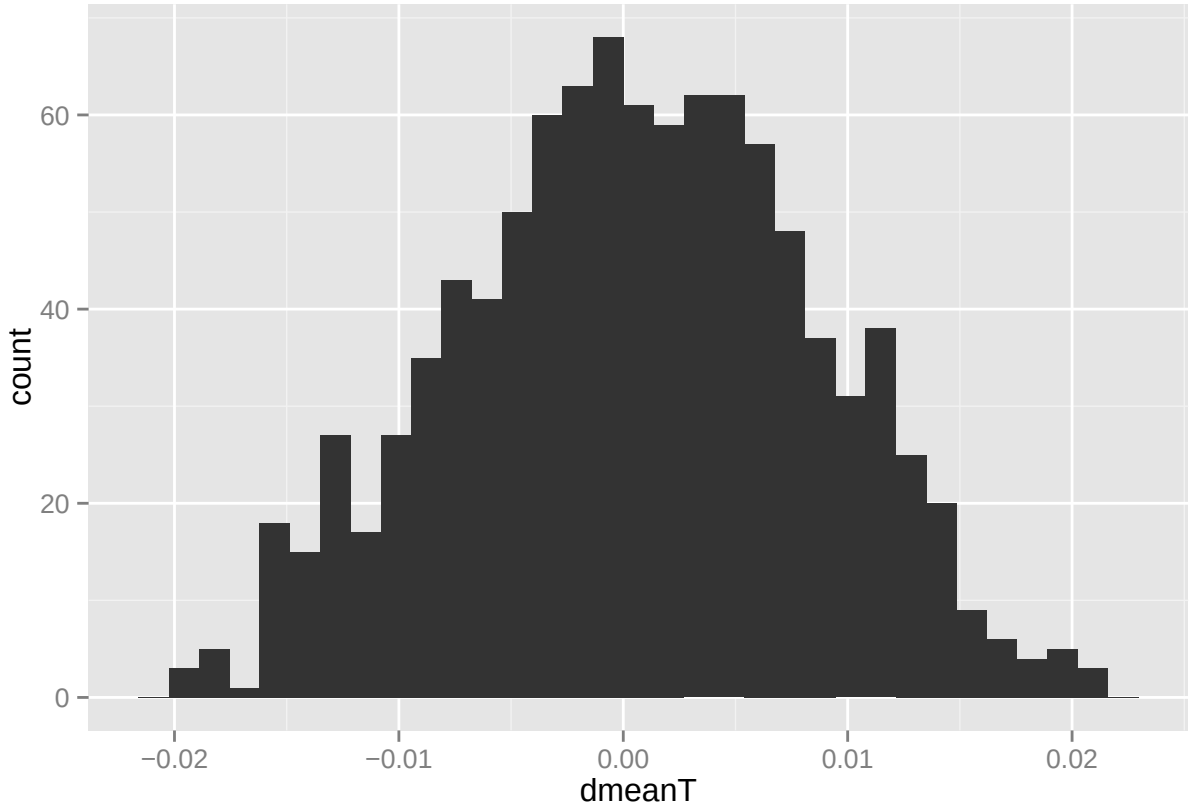
```
##
##  Welch Two Sample t-test
##
## data:  cntl and trt
## t = 3.7036, df = 11.671, p-value = 0.003156
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.009088475 0.035261525
## sample estimates:
## mean of x mean of y
##  0.231875  0.209700
```

```r
# really??
cntl.n <- length(cntl)
trt.n <- length(trt)
cntl.m <- mean(cntl)
trt.m <- mean(trt)
cntl.v <- var(cntl)
trt.v <- var(trt)

T <- (trt.m - cntl.m)/sqrt((trt.v/trt.n) + (cntl.v/cntl.n))
T
```

```
## [1] -3.703554
```

```r
# permutation test

samp <- c(cntl,trt)
l <- length(samp)
trt.n <- length(trt)

dmeanT <- NULL

for(i in 1:1000){

  sampler = sample((1:l),trt.n,replace=FALSE)
  trt.s = samp[sampler]
  cntl.s = samp[-sampler]
  dmean.s = mean(trt.s)-mean(cntl.s)
  dmeanT[i] = dmean.s
}

qplot(dmeanT)
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```




```
f <- ecdf(dmeanT)
quantile(dmeanT,.95)
```

```
##      95%
## 0.013375
```

```
quantile(dmeanT,.05)
```

```
##        5%
## -0.013175
```

```
summary(dmeanT)
```

```
##       Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
## -0.0192500 -0.0048500  0.0005500  0.0005538  0.0061750  0.0212500
```

```
# observed difference
permutation.t = cntl.m - trt.m
which(dmeanT>permutation.t) #0
```

```
## integer(0)
```

```
#p-value=0 < 0.05
#we can reject H0,
#there is enough evidence to reject the null hypothesis of no difference,
#and conclude that the ten essays were not actually written by Mark Twain.
```