Math 3070 § 1. Treibergs | **Soporific Example: Simulating the Sampling Distribution of the p-Value.** | Name: Example | June 21, 2011

One of the new topics added to our text in this edition, Devore's "Probability & Statistics for Engineering and the Sciences, 8th ed." is a discussion of the sampling distribution of the p-value, in section 8.4. The main point is that the p-value is a random variable. By simulating many t-tests, we can plot the histogram to appreciate the sampling distribution of the p-value.

Assume one selects random sample $X_1, X_2, \ldots, X_n \sim \mathrm{N}(\mu, \sigma)$ from a normal distribution. To test the hypothesis $H_0 : \mu = \mu_0$ vs. the alternative $H_a : \mu > \mu_0$, one computes the T statistic,

$$T = \frac{\bar{X} - \mu_0}{\frac{n}{\sqrt{n}}}.$$

which is also a random variable which is distributed according to the t-distribution with $n - 1$ degrees of freedom. In particular, any function of this is also a random variable, for example, the p-value

$$P = 1 - F(T) = \mathtt{pt(T, n-1, lower.tail = FALSE)},$$

where $F(x) = \mathrm{P}(T \le x)$ is the cdf for T with $n - 1$ degrees of freedom. If the background distribution has $\mu = \mu_0$, then the type I errors occur when P is small. The probability of a type I error is $\mathrm{P}(P \le \alpha)$ for a significance level α test, namely, that the test shows that the mean is significantly above μ_0 (*i.e.,* we reject H_0), even though the sample was drawn from data satisfying the null hypothesis $X_i \sim \mathrm{N}(\mu_0, \sigma)$.

It turns out that the p-value is a uniform rv in $[0, 1]$ when $\mu = \mu_0$. This is simply a consequence of definitions. Indeed, the cdf for P is

$$\begin{aligned}
\mathrm{P}(P \le \alpha) &= \mathrm{P}(1 - F(T) \le \alpha) \\
&= \mathrm{P}(F(T) \ge 1 - \alpha) \\
&= \mathrm{P}\left(T \ge F^{-1}(1 - \alpha)\right) \\
&= 1 - \mathrm{P}\left(T < F^{-1}(1 - \alpha)\right) \\
&= 1 - F\left(F^{-1}(1 - \alpha)\right) \\
&= 1 - (1 - \alpha) \\
&= \alpha,
\end{aligned}$$

which is the cdf for a uniform rv and so $P \sim \mathrm{U}(0, 1)$.

I ran examples with $\mu_0 = 0$, $\sigma = 1$, samples of size $n = 10$ and $m = 10{,}000$ trials for various $\mu = \mu_1$'s. In our histograms the bar from 0 to .05 is drawn red. For example, when $\mu_1 = \mu_0$, and so $P \sim \mathrm{U}(0, 1)$, the bars have nearly the same height and type I errors occurred 515 times or 5.15% of the time. If $\mu_1 > 0$, then the test is more likely to show that the mean is significantly greater than μ_0. As μ_1 increases, then the test is more and more likely to indicate that $\mu > \mu_0$ significantly. Note that if $\mu_1 < 0$, then the test is even less likely to make a type I error.

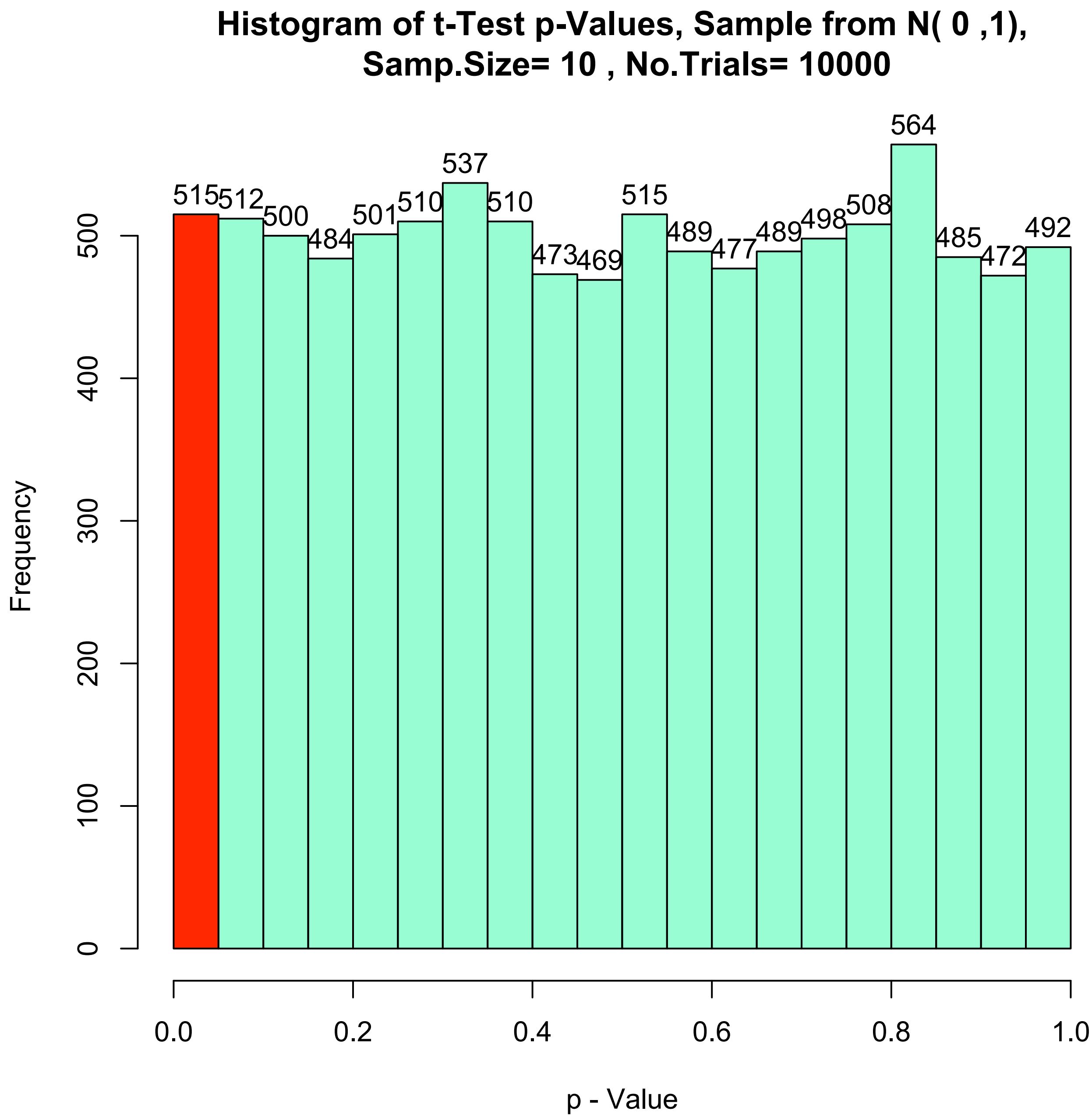
We start our **R** study by recalling Student's actual 1908 data to demonstrate his one-tailed t-test.

R Session:

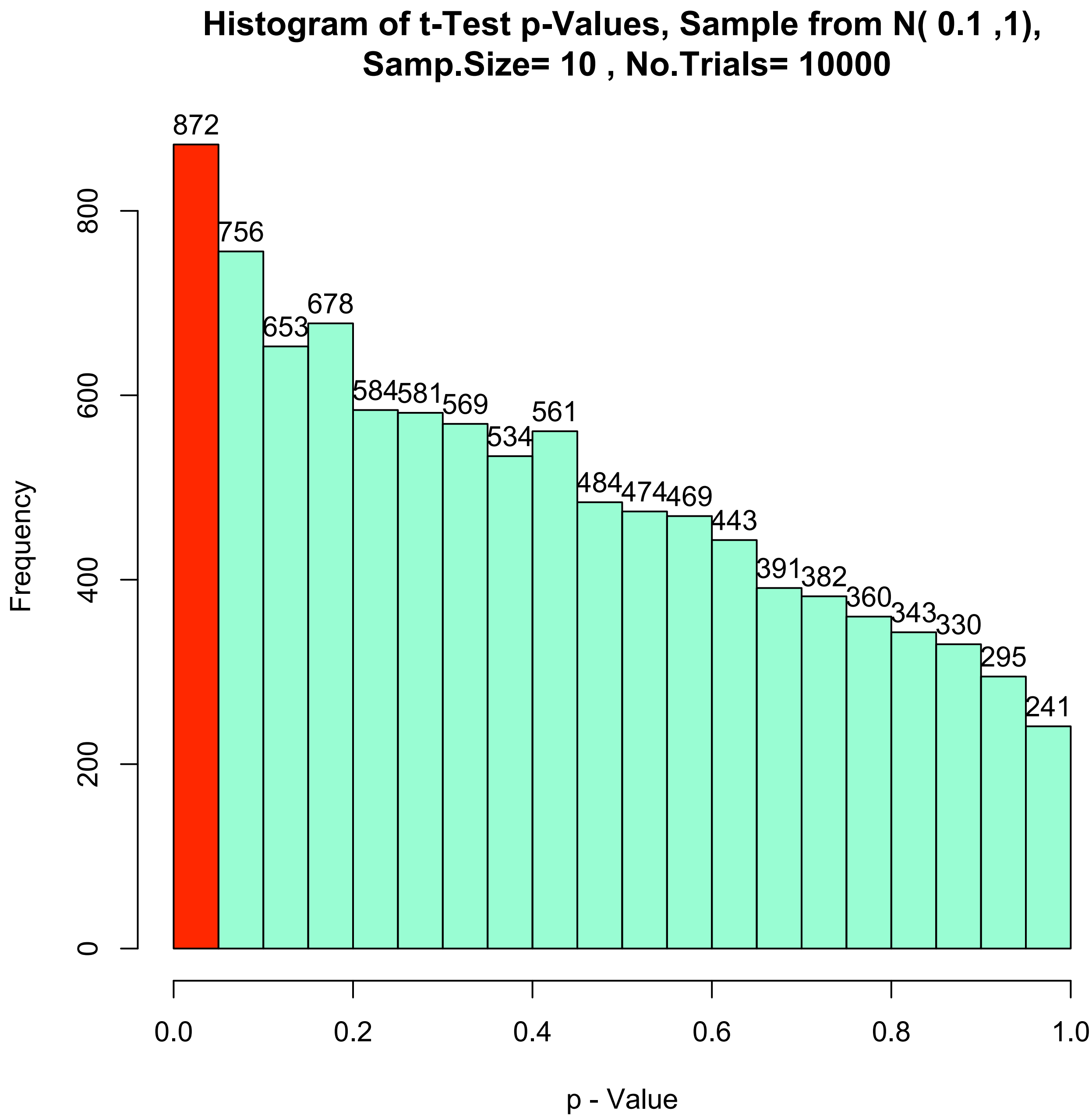
```
R version 2.10.1 (2009-12-14)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

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 1.31 (5538) powerpc-apple-darwin8.11.1]

[Workspace restored from /Users/andrejstreibergs/.RData]

> ############ STUDENT'S ONE-DAMPLE T-TEST DATA #########################
> # From M.G.Bulmer, "Principles of Statistics," Dover, 1979.
> # Student's 1908 Data:
> # Additional hrs sleep gained after administering Hyoscene
> # to ten patients
> #
> x <- scan()
1: 1.9 .8 1.1 .1 -.1 4.4 5.5 1.6 4.6 3.4
11:
Read 10 items
> x
 [1]  1.9  0.8  1.1  0.1 -0.1  4.4  5.5  1.6  4.6  3.4

> t.test(x,alternative="greater")

One Sample t-test

data:  x
t = 3.6799, df = 9, p-value = 0.002538
alternative hypothesis: true mean is greater than 0
95 percent confidence interval:
 1.169334      Inf
sample estimates:
mean of x
     2.33

> # Strong evidence that mu>0: Hyocene is soporific
```

```
> ################## ANALYZE THE DATA BY HAND ############################
> xbar <- mean(x); xbar
[1] 2.33
> s <- sd(x); s
[1] 2.002249
> n <- length(x); n
[1] 10
> t <- xbar/(s/sqrt(n)); t
[1] 3.679916
> # crit value for upper tailed test
> alpha <- .05
> qalpha <- qt(alpha, df = n-1, lower.tail = FALSE); qalpha
[1] 1.833113
> # T exceeds this so at 1-alpha conf., mu signif. greater than 0
> PV <- pt(t, n-1, lower.tail = FALSE); PV
[1] 0.002538066
> # same numbers as from canned test.

> ##################### SAMPLING DISTRIBUTION OF P-VALUE ##################
> # simulate p-values.
> # mu1 = mean of normal variable,  1 = sd of normal variable.
> # m = number of trials
> m <- 10000
> # n = sample size
> n <- 10
> # to make sure these computations are done outside the loop.
> c <- sqrt(n)
> nu <- n-1
> # Vector of nice colors.
> cl <- c(2,rep(rainbow(12,alpha=.4)[6],19))
> xl <- "p - Value"
>
> # The sapply(v,f)  does the function "f" to each element of vector "v"
> # In our case, f generates a p-value from a random sample every time
> # it's called.
> # In a vector oriented language, vectorwise computations replace loops
> # and do it faster.
>
> mu1 <- 0
> # Save the title.
> mn <- paste("Histogram of t-Test p-Values, Sample from N(",mu1,",1),
+ \n Samp.Size=", n,", No.Trials=", m)
> # Each call to  pv  gives p-value for a simulated upper-tail t-test
> # for size n.  Z  is a random  N(mu1,sigma)  sample of size  n
> pv <- function(j){Z<- rnorm(n,mu1,1);pt(c*mean(Z)/sd(Z),nu,lower.tail=F)}
> hist(sapply(1:m,pv),breaks=20,col=cl,labels=T,main=mn,xlab=xl)
> # M3074RVpValue1.pdf
```

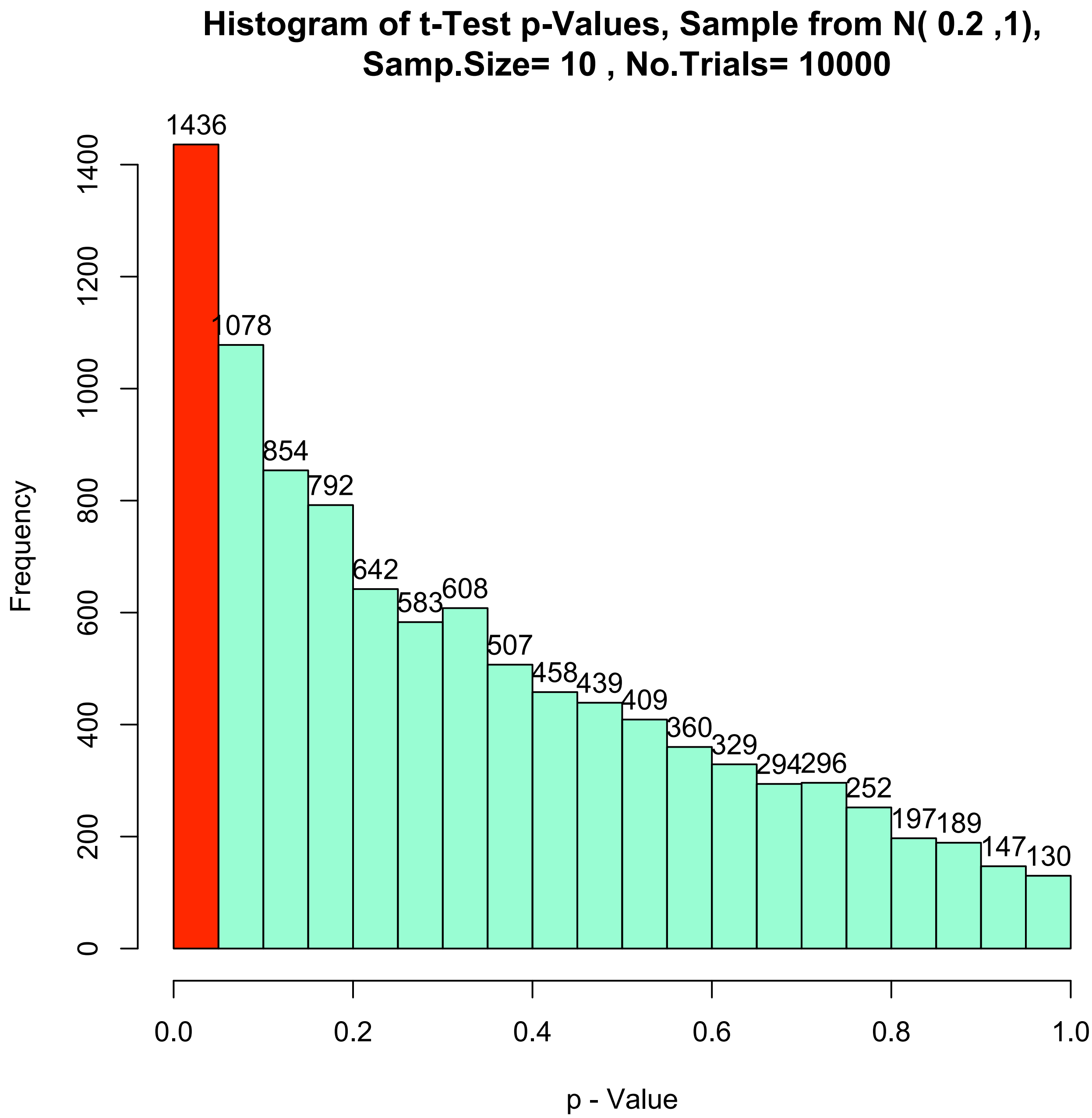
```
> mu1 <- 0.1
> mn <- paste("Histogram of t-Test p-Values, Sample from N(",mu1,",1),
+ \n Samp.Size=", n, ", No.Trials=", m)
> pv <- function(j){Z<- rnorm(n,mu1,1);pt(c*mean(Z)/sd(Z),nu,lower.tail=F)}
> hist(sapply(1:m,pv),breaks=20,col=cl,labels=T,main=mn,xlab=xl)
> # M3074RVpValue2.pdf
>
>
> mu1 <- 0.2
> mn <- paste("Histogram of t-Test p-Values, Sample from N(",mu1,",1),
+ \n Samp.Size=", n, ", No.Trials=", m)
> pv <- function(j){Z<- rnorm(n,mu1,1);pt(c*mean(Z)/sd(Z),nu,lower.tail=F)}
> hist(sapply(1:m,pv),breaks=20,col=cl,labels=T,main=mn,xlab=xl)
> # M3074RVpValue3.pdf
>
>
> mu1 <- 0.5
> mn <- paste("Histogram of t-Test p-Values, Sample from N(",mu1,",1),
+ \n Samp.Size=", n, ", No.Trials=", m)
> pv <- function(j){Z<- rnorm(n,mu1,1);pt(c*mean(Z)/sd(Z),nu,lower.tail=F)}
> hist(sapply(1:m,pv),breaks=20,col=cl,labels=T,main=mn,xlab=xl)
> # M3074RVpValue4.pdf
>
>
> mu1 <- 1
> mn <- paste("Histogram of t-Test p-Values, Sample from N(",mu1,",1),
+ \n Samp.Size=", n,", No.Trials=",m)
> pv <- function(j){Z<- rnorm(n,mu1,1);pt(c*mean(Z)/sd(Z),nu,lower.tail=F)}
> hist(sapply(1:m,pv),breaks=20,col=cl,labels=T,main=mn,xlab=xl)
> # M3074RVpValue5.pdf
>
>
> mu1 <- -.5
> mn <- paste("Histogram of t-Test p-Values, Sample from N(",mu1,",1),
+ \n Samp.Size=", n, ", No.Trials=", m)
> hist(sapply(1:m,pv),breaks=20,col=cl,labels=T,main=mn,xlab=xl)
> # M3074RVpValue6.pdf
```



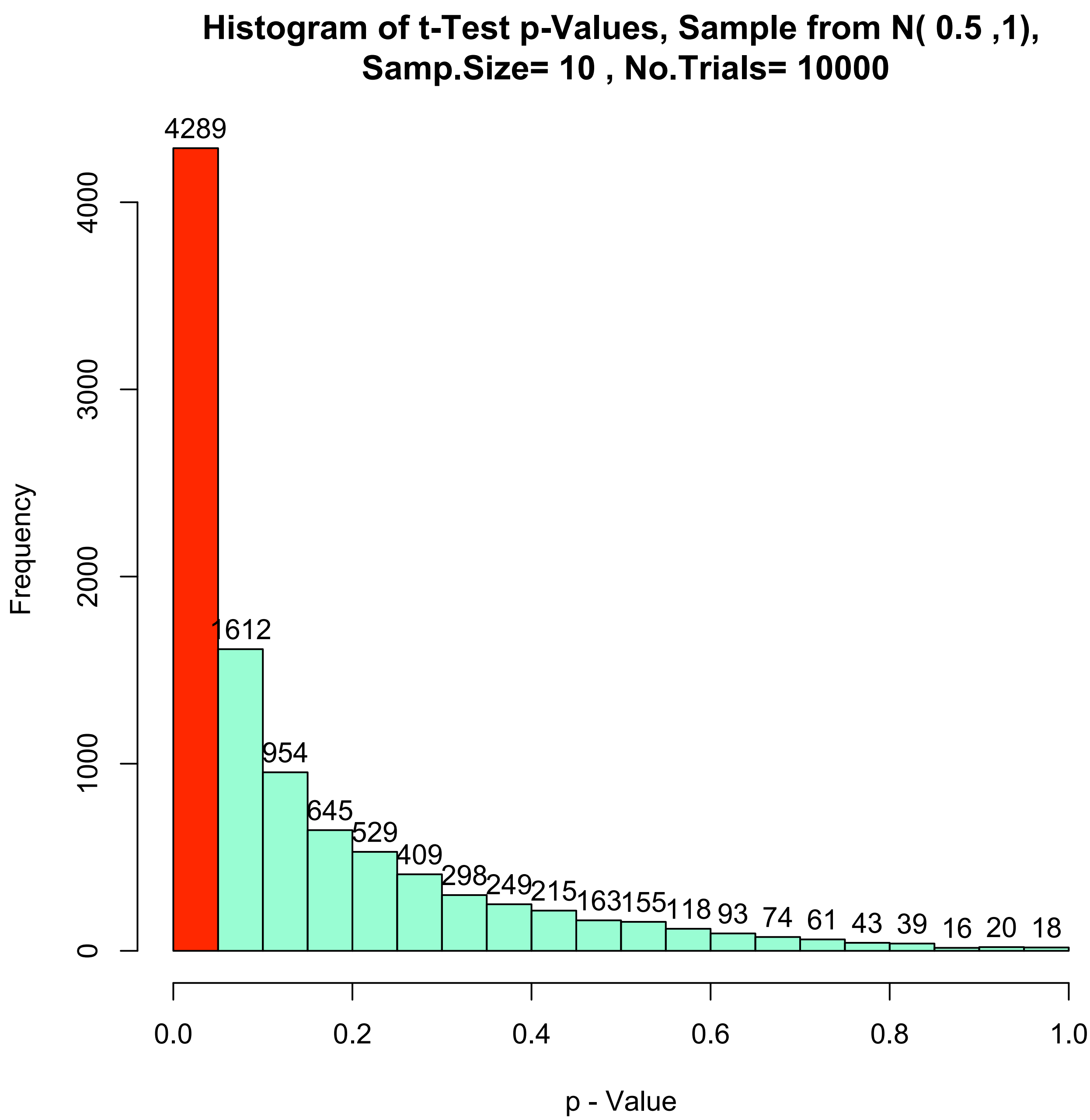
Histogram of t-Test p-Values, Sample from N(0 ,1),
Samp.Size= 10 , No.Trials= 10000
515
512
500
484
501
510
537
510
473
469
515
489
477
489
498
508
564
485
472
492
0
100
200
300
400
500
Frequency
0.0
0.2
0.4
0.6
0.8
1.0
p - Value

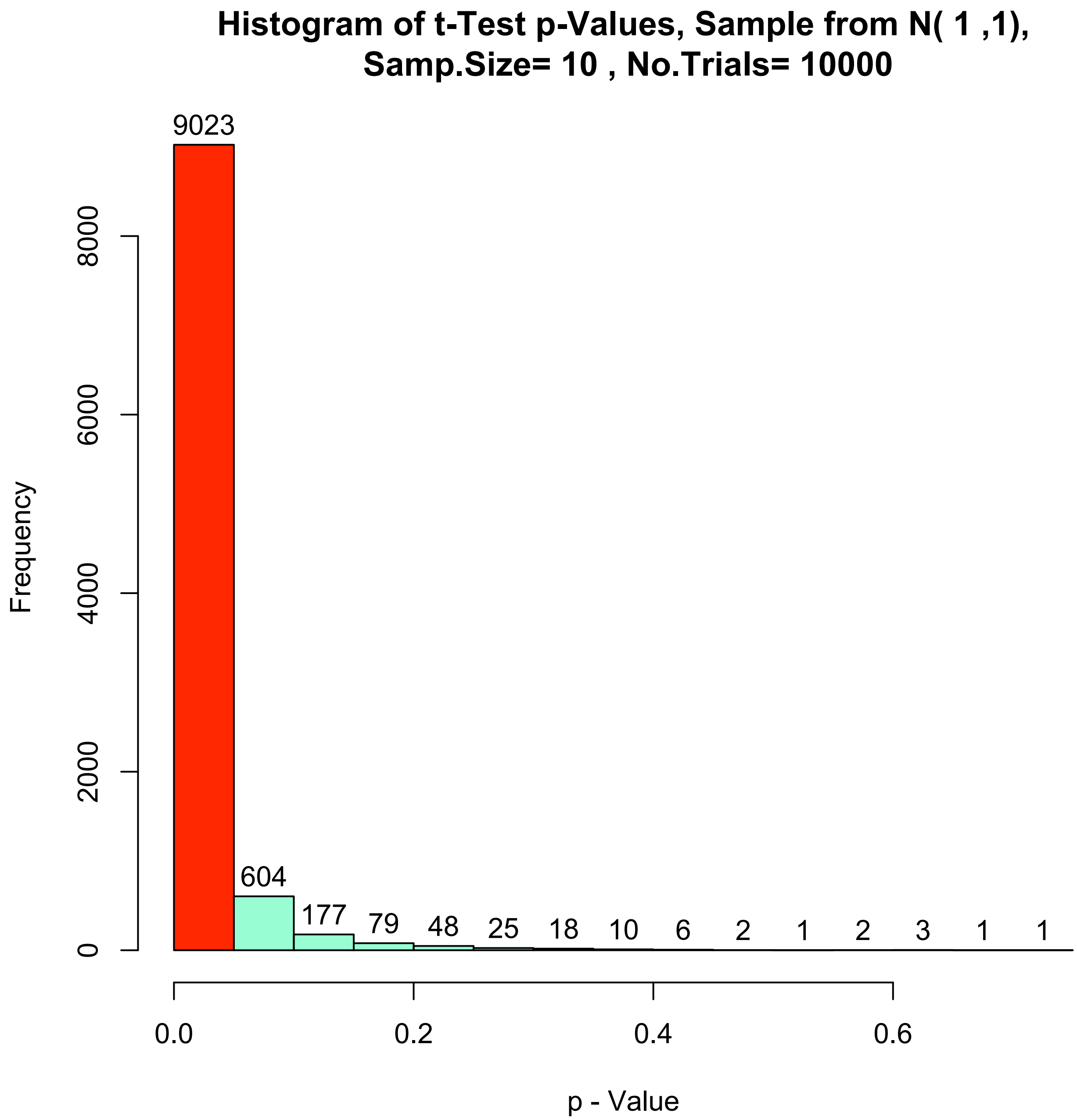


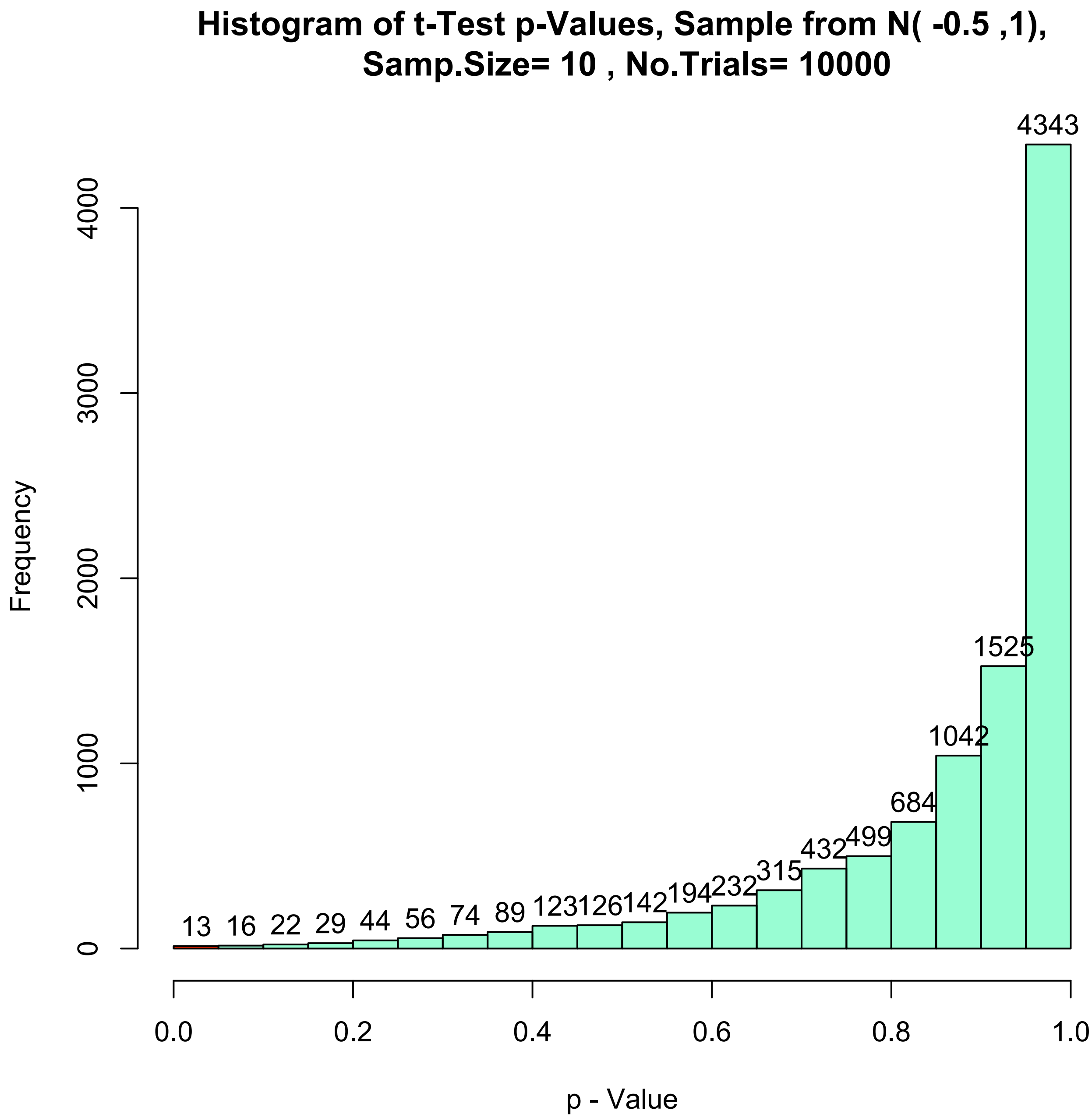
Histogram of t-Test p-Values, Sample from N(0.1 ,1),
Samp.Size= 10 , No.Trials= 10000
Frequency
p - Value
0
200
400
600
800
0.0
0.2
0.4
0.6
0.8
1.0
872
756
653
678
584
581
569
534
561
484
474
469
443
391
382
360
343
330
295
241



Histogram of t-Test p-Values, Sample from N(0.2 ,1), Samp.Size= 10 , No.Trials= 10000
Frequency
p - Value
0
200
400
600
800
1000
1200
1400
0.0
0.2
0.4
0.6
0.8
1.0
1436
1078
854
792
642
583
608
507
458
439
409
360
329
294
296
252
197
189
147
130


Histogram of t-Test p-Values, Sample from N(0.5 ,1),
Samp.Size= 10 , No.Trials= 10000
Frequency
p - Value
0
1000
2000
3000
4000
0.0
0.2
0.4
0.6
0.8
1.0
4289
1612
954
645
529
409
298
249
215
163
155
118
93
74
61
43
39
16
20
18


Histogram of t-Test p-Values, Sample from N(1 ,1),
Samp.Size= 10 , No.Trials= 10000
9023
604
177
79
48
25
18
10
6
2
1
2
3
1
1
Frequency
0
2000
4000
6000
8000
0.0
0.2
0.4
0.6
p - Value


Histogram of t-Test p-Values, Sample from N(-0.5 ,1),
Samp.Size= 10 , No.Trials= 10000
Frequency
0
1000
2000
3000
4000
13 16 22 29 44 56 74 89 123 126 142 194 232 315 432 499 684 1042 1525 4343
0.0
0.2
0.4
0.6
0.8
1.0
p - Value