

# Stat 3202 Lab 10

Jane Weissberg.11

2024-07-25

```
library(DescTools)
```

1a.

```
oranges1 <- c(126,122,129,128,133,128,119,124,134,
134,128,125,127,125,125,126,131,125,
124,124,133,126,132,131,128,124,131,
130,120,131,126,130,121,118,129,123)

alpha <- 0.05
mu_null <- 125
mu_reality <- 125

t.test(oranges1, alternative="two.sided", mu=125, conf.level=1-alpha)
```

```
##
## One Sample t-test
##
## data: oranges1
## t = 2.7812, df = 35, p-value = 0.008663
## alternative hypothesis: true mean is not equal to 125
## 95 percent confidence interval:
## 125.5251 128.3638
## sample estimates:
## mean of x
## 126.9444
```

1b.

```
set.seed(515)
n <- 36
alpha <- 0.05
mu_null <- 125
mu_reality <- 125
sigma <- 5

p_storage <- c()

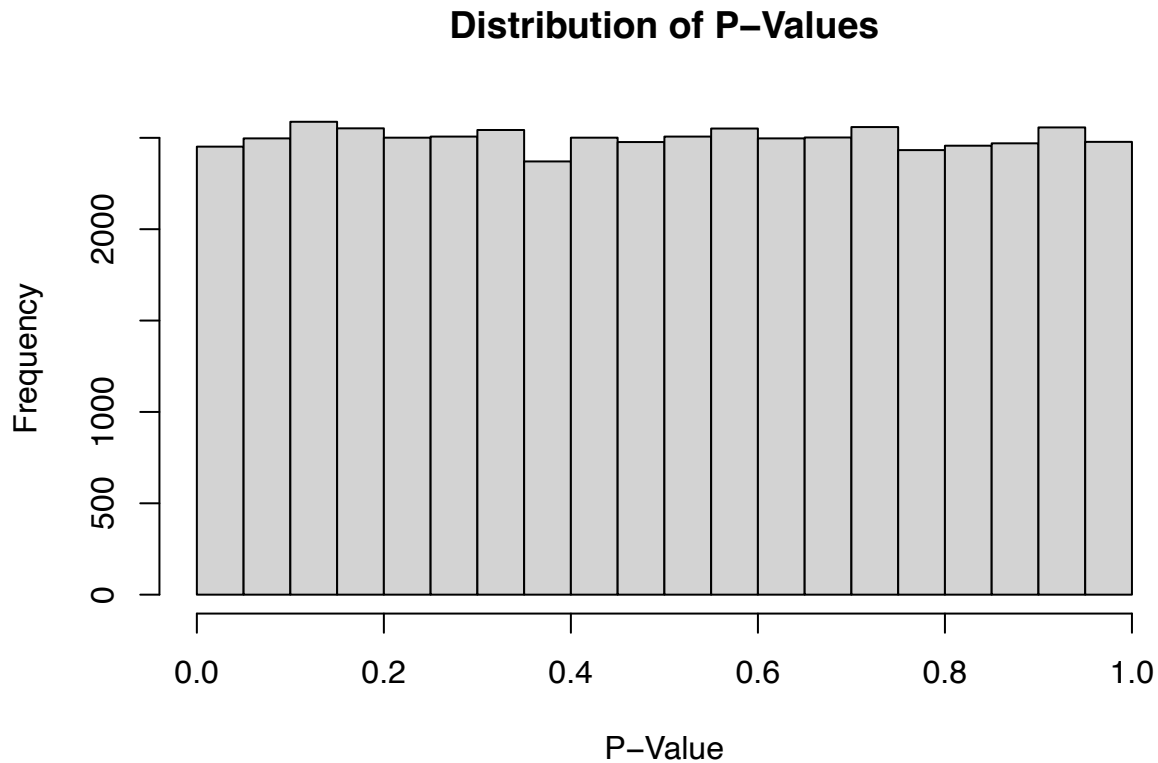
for (i in 1:50000) {
  x <- rnorm(n, mean=mu_null, sd=sigma)
```

```

    out <- t.test(x, alternative="two.sided", mu=125, conf.level=1-alpha)
    p_storage[i] <- out$p.value
  }

hist(p_storage, main = "Distribution of P-Values", xlab = "P-Value")

```



1c. The histogram of the stored p-values from the Monte Carlo Simulation is approximately uniform. This did surprise me because I was not sure what to expect. I thought there would be some p-values that had higher frequency than others.

1d (alpha=0.05).

```

set.seed(515)
n <- 36
alpha <- 0.05
mu_null <- 125
mu_reality <- 125
sigma <- 5

results <- c()

for (i in 1:50000) {
  x <- rnorm(n, mean=mu_null, sd=sigma)
  out <- t.test(x, alternative="two.sided", mu=125, conf.level=1-alpha)
  results[i] <- out$p.value < alpha
}

```

```
mean(results)
```

```
## [1] 0.04904
```

A proportion of approximately 0.05 of my results led to a rejection of the null.  
1d (alpha=0.01).

```
set.seed(515)
n <- 36
alpha <- 0.01
mu_null <- 125
mu_reality <- 125
sigma <- 5

p_storage <- c()

for (i in 1:50000) {
  x <- rnorm(n, mean=mu_null, sd=sigma)
  out <- t.test(x, alternative="two.sided", mu=125, conf.level=1-alpha)
  results[i] <- out$p.value < alpha
}

mean(results)
```

```
## [1] 0.0099
```

A proportion of approximately 0.01 of my results led to a rejection of the null.  
1d (alpha=0.84).

```
set.seed(515)
n <- 36
alpha <- 0.84
mu_null <- 125
mu_reality <- 125
sigma <- 5

p_storage <- c()

for (i in 1:50000) {
  x <- rnorm(n, mean=mu_null, sd=sigma)
  out <- t.test(x, alternative="two.sided", mu=125, conf.level=1-alpha)
  results[i] <- out$p.value < alpha
}

mean(results)
```

```
## [1] 0.83912
```

A proportion of approximately 0.84 of my results led to a rejection of the null.

1e. The type-I error rate when the null is true is the significance level, or  $\alpha$ . This relates to what we have been working on with rejection regions and p-values because the significance level corresponds to the values that lead to rejection regions and p-values.

2a.

```
set.seed(515)
n <- 31
mu <- 16
variance <- 9
alpha <- 0.01

results <- c()
crit_value <- qchisq(1-alpha, df=n-1)

for (i in 1:50000) {
  x <- rnorm(n, mean=mu, sd=sqrt(variance))
  sample_var <- var(x)
  test_stat <- (n-1)*sample_var/variance
  results[i] <- test_stat>crit_value
}

mean(results)
```

```
## [1] 0.00934
```

You can lower the type-I error rate by lowering the significance level or increasing the sample size.

3.

```
set.seed(515)

n <- 31
mu <- 16
var_null <- 4
var_reality <- 9
alpha <- 0.01

results <- c()

for (i in 1:50000) {
  x <- rnorm(n, mu, sqrt(var_reality))
  out <- VarTest(x, alternative="greater", sigma.squared = var_null, conf.level=1-alpha)
  results[i] <- out$p.value>alpha
}

mean(results)
```

```
## [1] 0.16948
```

The Type-II error rate is 0.17. This tells us that there is 0.17 probability that we will fail to reject the null hypothesis, even though the null hypothesis is false.