

## 562\_HW\_3

Matthew Stoebe

2024-11-10

#Question 1: In Lecture 39, we discussed an example related to soil contamination (slide 14). The data are: {6.2, 4.1, 3.5, 5.1, 5.0, 3.6, 4.8, 4.1, 3.6, 4.7, 4.3, 4.2} We performed a sign test to evaluate the hypothesis that the median was 4.0 ppm. Perform a one-sample t-test to test  $H_0 : \mu = 4$  vs  $H_a : \mu > 4$

```
soil_data <- c(6.2, 4.1, 3.5, 5.1, 5.0, 3.6, 4.8, 4.1, 3.6, 4.7, 4.3, 4.2)
```

##a State the test statistic and the distribution of the test statistic under  $H_0$ .

The test statistic is a t test which under the null hypothesis follows the t distribution with 11 (n-1) degrees of freedom

##b Compute the observed test statistic and the associated p-value.

```
# Perform one-sample t-test
t_test_result <- t.test(soil_data, mu = 4, alternative = "greater")

# Display the results
t_test_result

##
## One Sample t-test
##
## data:  soil_data
## t = 1.9311, df = 11, p-value = 0.03982
## alternative hypothesis: true mean is greater than 4
## 95 percent confidence interval:
##  4.030344      Inf
## sample estimates:
## mean of x
##  4.433333
```

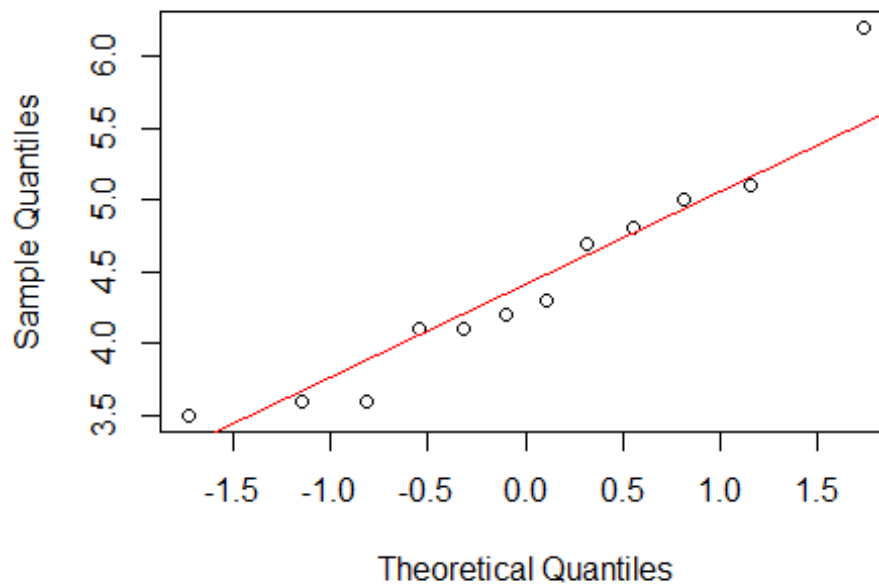
##c

Create a probability plot for the data to assess whether it appears normally distributed. Based on your plot do you believe the results of the sign test or t-test more?

The data seems relatively normal. I trust the Sign test more because I know it is less likely to incur type 1 error, but I think because the Data appears normal and the spreads were not super large, I would reject the null hypothesis with the T test with relative confidence.

```
qqnorm(soil_data, main = "Q-Q Plot of Soil Contamination Data")
qqline(soil_data, col = "red")
```

### Q-Q Plot of Soil Contamination Data



#Question 2 Recall

that  $\alpha$  represents the probability of making a Type I error (i.e. reject  $H_0$  when  $H_0$  is true). When the distribution of our test statistic is continuous (i.e. z-test, t-test, chi-squared test), we can determine a decision rule with Type I error probability equal to  $\alpha = .05$  exactly. However, for the sign test, our test statistic has a Binomial distribution, which is discrete. Thus, in most cases we will not be able to find a decision rule with the desired  $\alpha$ . Instead, we will devise a rule that has level as close to  $\alpha$  as possible, without exceeding  $\alpha$ . Consider a sign test with 5 observations where we are interested in testing  $H_0 : m = 10$  versus  $H_a : m < 10$ . Let  $X$  denote the number of observations less than 10.

##a

```
n <- 5
p <- 0.5
alpha <- 0.20

cum_probs <- pbinom(0:n, size = n, prob = p)

# Find the critical value
x_critical <- max(which(cum_probs <= alpha))

# Actual Level
actual_level <- pbinom(x_critical, size = n, prob = p)

cat("Critical value for alpha = 0.20:", x_critical, "\n")

## Critical value for alpha = 0.20: 2
```

```

cat("Actual level:", actual_level, "\n")
## Actual level: 0.5

##b

alpha <- 0.10

# Find the critical value
x_critical <- max(which(cum_probs <= alpha))
# Actual Level
actual_level <- pbinom(x_critical, size = n, prob = p)

cat("Critical value for alpha = 0.10:", x_critical, "\n")
## Critical value for alpha = 0.10: 1

cat("Actual level:", actual_level, "\n")
## Actual level: 0.1875

##c

alpha <- 0.05

# Find the critical value
x_critical <- max(which(cum_probs <= alpha))

# Actual Level
actual_level <- pbinom(x_critical, size = n, prob = p)

cat("Critical value for alpha = 0.05:", x_critical, "\n")
## Critical value for alpha = 0.05: 1

cat("Actual level:", actual_level, "\n")
## Actual level: 0.1875

```

#Question 3 A national highway safety association is interested in measuring the difference in reaction times for drivers when they have had just one drink compared to when they are completely sober. Reaction times to avoid a simulated car crash are measured using a driving simulation machine. There are fifteen volunteers with ages ranging from 21-65, with both genders represented. A paired-samples design is chosen. The test is administered twice to each subject, with testing times a week apart. On one occasion, the subject has had one serving of alcohol administered 30 minutes before the test, and on the other occasion the subject has had no alcohol. A coin is tossed to determine which treatment is administered first.

##a Define Parameters and State the Hypotheses

$d_i = \text{drunk\_reaction\_time}_i - \text{sober\_reaction\_time}_i$   $\mu_d = \text{mean of the differences}$   $d$

$H_0 = \mu_d = 0$   $H_A = \mu_d > 0$

This is a one tail test

##b Explain why a paired-samples design is chosen instead of an independent samples design. Here are the reaction times (hundreths of seconds)

A paired sample test is best in this case as it controls for the individual baseline reaction speeds of each driver. It also will most likely have greater power.

##c

```
drink <- c(111.3, 108.0, 113.7, 114.6, 112.3, 111.8, 109.8, 114.1, 111.7,
111.9, 113.4, 113.1, 117.1, 113.7, 119.9)
sober <- c(109.8, 112.0, 113.7, 111.2, 110.9, 112.0, 106.3, 113.5, 105.9,
114.5, 111.6, 111.8, 110.2, 108.7, 112.4)

paired_t_test <- t.test(drink, sober, alternative = "greater", paired = TRUE)
paired_t_test

##
## Paired t-test
##
## data: drink and sober
## t = 2.5099, df = 14, p-value = 0.01249
## alternative hypothesis: true mean difference is greater than 0
## 95 percent confidence interval:
## 0.6342754 Inf
## sample estimates:
## mean difference
## 2.126667
```

#Question 4 To determine if the economic health in Larimer County has improved the three years receiving stimulus funds, the change in net worth was measured for a random sample of residents. A negative measure means the net worth has decreased. The measurements (in \$1000) are {-7.1, 240.6, 57.0, 36.8, 2.3, 42.2, 126.8, 11.4, 184.2, -35.1, -24.2, -26.3}.

##a Do a sign test and report the conclusion in the context of the problem with  $\alpha = 0.075$ .

```
net_worth <- c(-7.1, 240.6, 57.0, 36.8, 2.3, 42.2, 126.8, 11.4, 184.2, -35.1,
-24.2, -26.3)

n_positive <- sum(net_worth > 0) # 8
n_negative <- sum(net_worth < 0) # 4

n <- n_positive + n_negative # 12

sign_test_result <- binom.test(n_positive, n, p = 0.5, alternative =
"greater")
```

```

print(sign_test_result)

##
## Exact binomial test
##
## data: n_positive and n
## number of successes = 8, number of trials = 12, p-value = 0.1938
## alternative hypothesis: true probability of success is greater than 0.5
## 95 percent confidence interval:
## 0.3908623 1.0000000
## sample estimates:
## probability of success
## 0.6666667

```

Because our P value is  $> .075$ , we are not able to reject our null hypothesis

##b Do a one-sample t-test of  $H_0 : \mu = 0$  versus  $H_a : \mu > 0$  where  $\mu$  is the mean change in net worth of Larimer County residents

```

t_test_net_worth <- t.test(net_worth, mu = 0, alternative = "greater")

print(t_test_net_worth)

##
## One Sample t-test
##
## data: net_worth
## t = 1.9857, df = 11, p-value = 0.03628
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
## 4.847993      Inf
## sample estimates:
## mean of x
## 50.71667

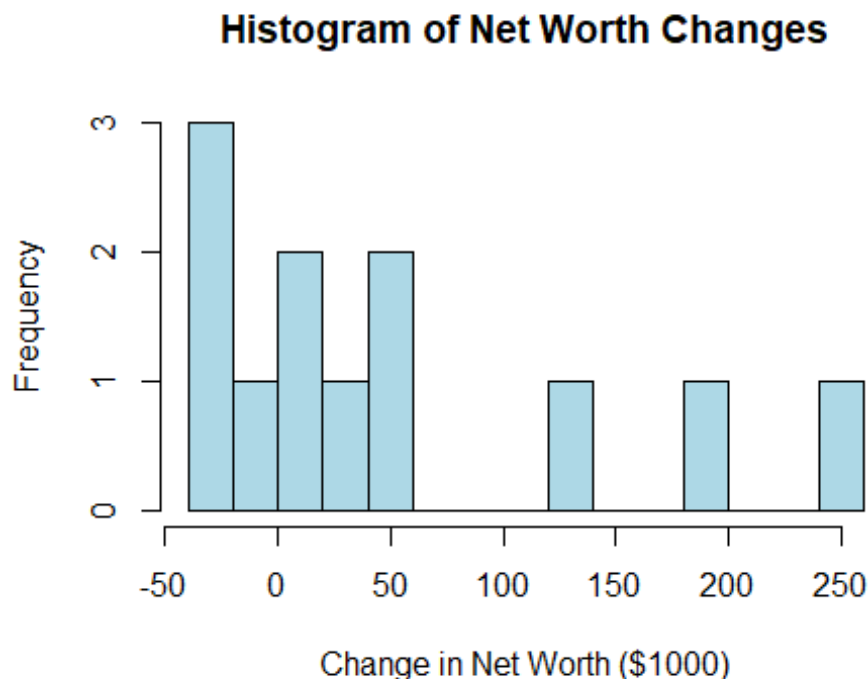
```

##c Get a histogram of the data and interpret what you see in the context of the problem. Do you think your answer in (a) or your answer in (b) is more accurate?

```

hist(net_worth, breaks = 10, col = "lightblue",
     main = "Histogram of Net Worth Changes",
     xlab = "Change in Net Worth ($1000)",
     border = "black")

```



This data does not appear normal, has relatively few data points and has high positive outliers. As such, it is not reasonable to use a T test making the non parametric sign test more accurate.

##d Compute the power for your test in part (a) if zero change in net worth represents the 30th percentile.

```
n <- 12
p_alternative <- 0.7 # 30th percentile
x_critical <- 8

power <- 1 - pbinom(x_critical - 1, size = n, prob = p_alternative)

cat("Power of the sign test when p = 0.7:", round(power, 4), "\n")

## Power of the sign test when p = 0.7: 0.7237
```

#Question 5 Let's visualize the CLT with simulation.

##a We know that if we take a sample of size  $n$  from a  $\text{Gamma}(2, \text{rate}=1/2)$  density, the sample mean is a random variable with mean 4 and variance  $8/n$ . We can take a look at the distribution of the sample mean by simulating 10,000 (or more) data sets of size 20 from this density and obtaining 10,000 sample means. Plot the histogram of the sample means using `freq=FALSE` and overlay a normal density with mean 4 and variance .4. Repeat for  $n = 100$ , and notice that the normal density fits better. (Note when you change the sample size or the distribution, you need to recalculate the appropriate normal approximation.) Below is code to get you started for  $n = 20$ :

```

sims <- 10000      # Number of simulations
n <- 20           # Sample size
shape <- 2        # Shape parameter for Gamma
rate <- 0.5       # Rate parameter for Gamma

# Initialize vector to store sample means
store_means <- numeric(sims)

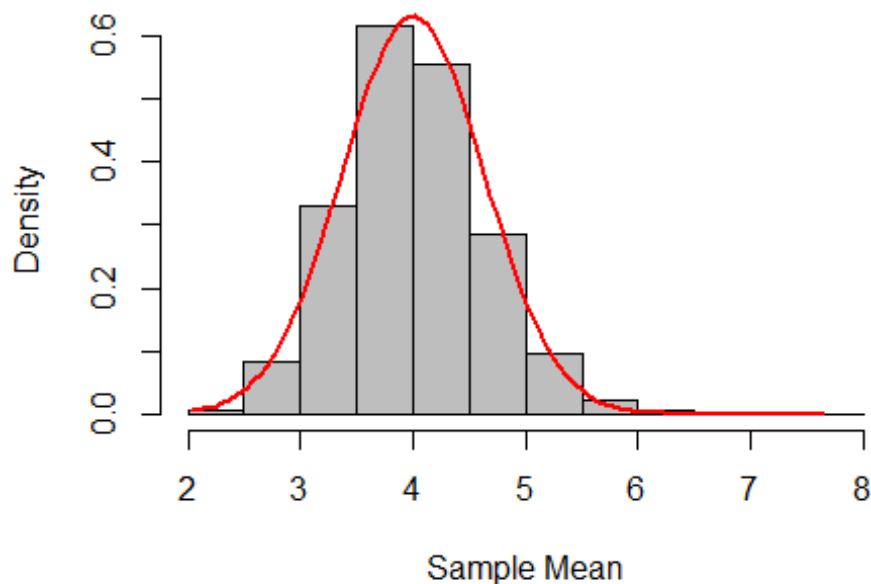
# Simulation loop
set.seed(123)     # For reproducibility
for(i in 1:sims){
  samp <- rgamma(n, shape=shape, rate=rate) # Generate sample
  store_means[i] <- mean(samp)             # Compute sample mean
}

# Plot histogram of sample means
hist(store_means, freq=FALSE, col="grey",
      main="Histogram of Sample Means (Gamma(2, 0.5), n=20)",
      xlab="Sample Mean",
      border = "black")

# Overlay normal density
x_range <- seq(min(store_means), max(store_means), length=100)
normal_density <- dnorm(x_range, mean=4, sd=sqrt(8/n)) # mean=4, sd=0.6325
lines(x_range, normal_density, col="red", lwd=2)

```

**Histogram of Sample Means (Gamma(2, 0.5), n=20)**



##b

```

sims <- 10000      # Number of simulations
n <- 20            # Sample size
p <- 0.25          # Probability of success

store_means_b <- numeric(sims)

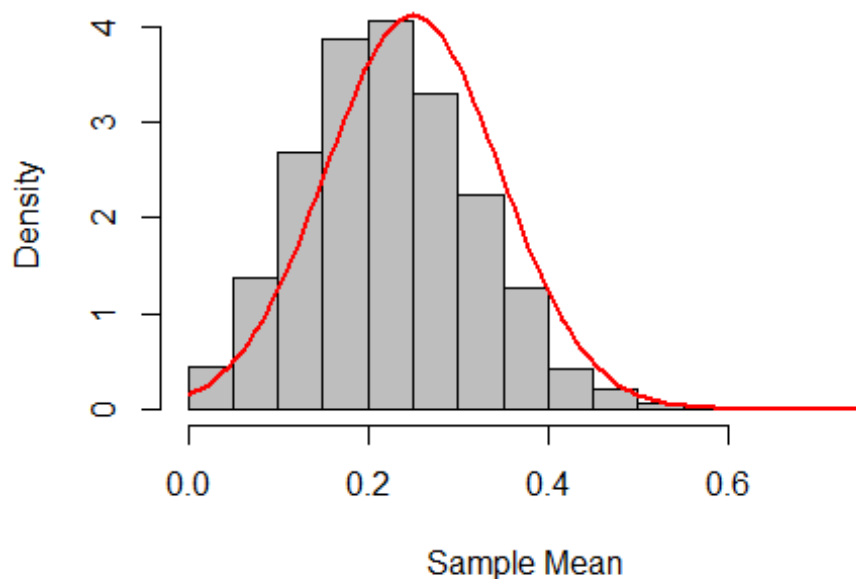
set.seed(123)
for(i in 1:sims){
  samp <- rbinom(n, size=1, prob=p) # Generate Bernoulli sample
  store_means_b[i] <- mean(samp)    # Compute sample mean
}

hist(store_means_b, freq=FALSE, col="grey",
      main="Histogram of Sample Means (Bernoulli(0.25), n=20)",
      xlab="Sample Mean",
      border = "black")

# Overlay normal density
x_range_b <- seq(min(store_means_b), max(store_means_b), length=100)
normal_density_b <- dnorm(x_range_b, mean=p, sd=sqrt(p*(1-p)/n)) #
# mean=0.25, sd≈0.0968
lines(x_range_b, normal_density_b, col="red", lwd=2)

```

## Histogram of Sample Means (Bernoulli(0.25), n=20)



##c

```

sims <- 10000      # Number of simulations
n <- 20            # Sample size
shape1 <- 2        # Alpha parameter for Beta

```



```

shape2 <- 2          # Beta parameter for Beta

store_means_c <- numeric(sims)

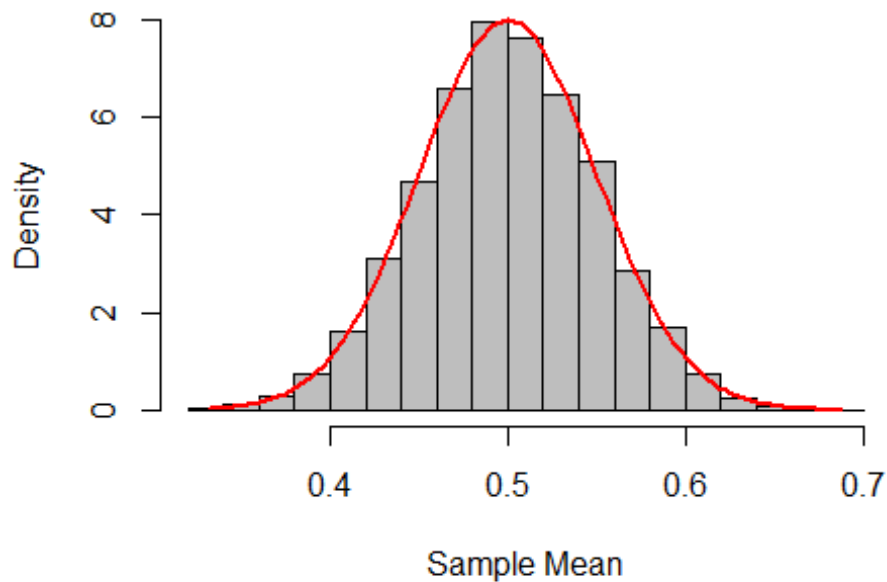
for(i in 1:sims){
  samp <- rbeta(n, shape1=shape1, shape2=shape2) # Generate Beta sample
  store_means_c[i] <- mean(samp)                 # Compute sample mean
}

hist(store_means_c, freq=FALSE, col="grey",
      main="Histogram of Sample Means (Beta(2,2), n=20)",
      xlab="Sample Mean",
      border = "black")

x_range_c <- seq(min(store_means_c), max(store_means_c), length=100)
normal_density_c <- dnorm(x_range_c, mean=0.5, sd=sqrt(0.05/n)) # mean=0.5,
sd=0.05
lines(x_range_c, normal_density_c, col="red", lwd=2)

```

**Histogram of Sample Means (Beta(2,2), n=20)**



#Question 6 A

weaving machine in a garment factory produces lengths of linen cloth. In a randomly selected yard of cloth, the number of flaws follows this distribution. flaws {0, 1, 2, 3} probability {.6, .2, .15, .05} The manager of the factory has the opportunity to buy a new machine and wants to test the null hypothesis that the distribution of numbers of flaws for the new machine is the same as for the old machine against the alternative that the new machine makes, on average, fewer flaws. The manager will make 20 yards of linen cloth on

the new machine. Assume that the numbers of flaws in each yard are independent random variables.

```
##a

flaws <- 0:3
probabilities <- c(0.6, 0.2, 0.15, 0.05)

# Calculate the population mean ( $\mu$ )
mu <- sum(flaws * probabilities)

EX2 <- sum((flaws^2) * probabilities)

variance <- EX2 - mu^2
sigma <- sqrt(variance)

cat("Population Mean:", mu, "\n")
## Population Mean: 0.65

cat("Population Variance:", variance, "\n")
## Population Variance: 0.8275

cat("Population Standard Deviation:", sigma, "\n\n")
## Population Standard Deviation: 0.9096703

n <- 20

sigma_Xbar <- sigma / sqrt(n)

# Significance Level
alpha <- 0.046

z_alpha <- qnorm(alpha)
c <- mu + z_alpha * sigma_Xbar

cat("Standard Deviation of Sample Mean ( $\sigma_{\bar{X}}$ ):", sigma_Xbar, "\n")
## Standard Deviation of Sample Mean ( $\sigma_{\bar{X}}$ ): 0.2034085

cat("Z-score for alpha =", alpha, ":", z_alpha, "\n")
## Z-score for alpha = 0.046 : -1.684941

cat("Critical Value (c):", c, "\n\n")
## Critical Value (c): 0.3072688
```

```
cat("Decision Rule: Reject  $H_0$  if the sample mean number of flaws <",  
    round(c, 3), "\n")
```

```
## Decision Rule: Reject  $H_0$  if the sample mean number of flaws < 0.307
```

##b Because the distribution of the test statistic is not really correct, we don't necessarily reject with probability .046 when the null hypothesis is true. Do simulations in R to find the true test size. (Note you can sample from the flaw distribution using the command `sample(1,x=c(0:3),prob=c(.6,.2,.15,.05))`)

```
# Parameters
```

```
n <- 20
```

```
alpha <- 0.046
```

```
c <- 0.311
```

```
sims <- 10000 # Number of simulations
```

```
set.seed(123) # For reproducibility
```

```
simulated_samples <- replicate(sims, sample(flaws, size = n, replace = TRUE,  
prob = probabilities))
```

```
sample_means <- colMeans(simulated_samples)
```

```
# Calculate the proportion of sample means less than the critical value 'c'  
true_test_size <- mean(sample_means < c)
```

```
# Display the true test size
```

```
cat("True Test Size (Proportion of Rejections when the null hypothesis is  
true is true):", round(true_test_size, 4), "\n")
```

```
## True Test Size (Proportion of Rejections when the null hypothesis is true  
is true): 0.046
```

##c Suppose the true distribution for flaws with the new machine is flaws {0, 1, 2, 3} probability {0.75, 0.15, 0.1, 0} Do simulations to determine the power for the test in (a)

```
# Define the new flaws distribution under  $H_a$ 
```

```
flaws_new <- 0:3
```

```
probabilities_new <- c(0.75, 0.15, 0.1, 0)
```

```
n <- 20
```

```
c <- 0.311
```

```
sims <- 100000
```

```
simulated_samples_new <- replicate(sims, sample(flaws_new, size = n, replace  
= TRUE, prob = probabilities_new))
```

```
# Compute the sample means
```

```
sample_means_new <- colMeans(simulated_samples_new)
```

```
power <- mean(sample_means_new < c)
```

```
cat("Power of the Test (Proportion of Rejections when Alternative Hypothesis is true):", round(power, 4), "\n")
```

```
## Power of the Test (Proportion of Rejections when Alternative Hypothesis is true): 0.4609
```

##d Write code to simulate the true distribution of T, the average number of flaws in the 20 yards, when the null hypothesis is true. What is the decision rule that gives a test size as close as possible to .05?

```
desired_alpha <- 0.05
```

```
sims <- 100000 # Number of simulations
```

```
simulated_samples_H0 <- replicate(sims, sample(flaws, size = n, replace = TRUE, prob = probabilities))
```

```
# Compute the sample means and Critical value
```

```
sample_means_H0 <- colMeans(simulated_samples_H0)
```

```
c_prime <- quantile(sample_means_H0, probs = desired_alpha)
```

```
actual_alpha <- mean(sample_means_H0 < c_prime)
```

```
cat("Critical Value (c') for Test Size  $\approx$  0.05:", round(c_prime, 3), "\n")
```

```
## Critical Value (c') for Test Size  $\approx$  0.05: 0.35
```

```
cat("Actual Test Size:", round(actual_alpha, 4), "\n")
```

```
## Actual Test Size: 0.0467
```

```
n <- 20
```

```
c_prime <- 0.313
```

```
sims <- 100000 # Number of simulations
```

```
simulated_samples_new_Ha <- replicate(sims, sample(flaws_new, size = n, replace = TRUE, prob = probabilities_new))
```

```
# Compute the sample means and power
```

```
sample_means_new_Ha <- colMeans(simulated_samples_new_Ha)
```

```
power_new <- mean(sample_means_new_Ha < c_prime)
```

```
cat("Power of the Test (Proportion of Rejections when Alternative Hypothesis is true):", round(power_new, 4), "\n")
```

```
## Power of the Test (Proportion of Rejections when Alternative Hypothesis is true): 0.4563
```