[1.0/1.2] total grade: 5/6

Q.1 a]

Formulating hypothesis :-

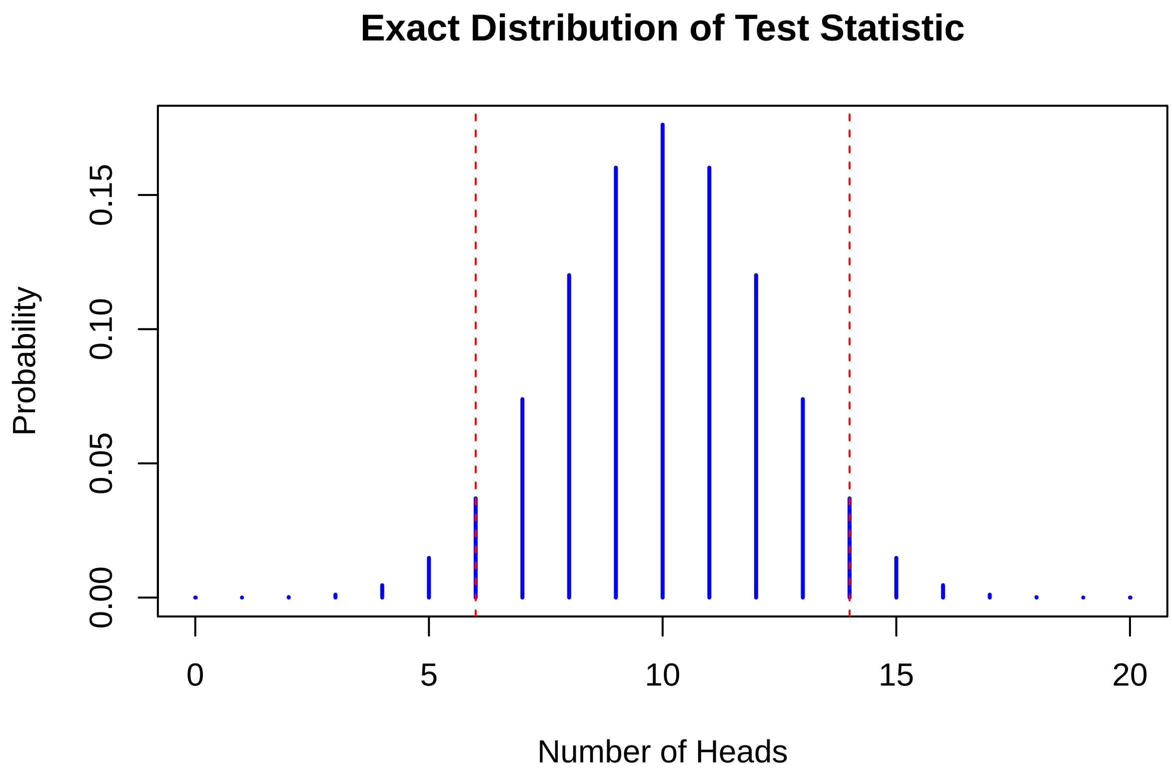
Null hypothesis :- Ho , The coin is fair & the probability of getting head (p) is equal to 0.5

i.e Ho : $p = 0.5$

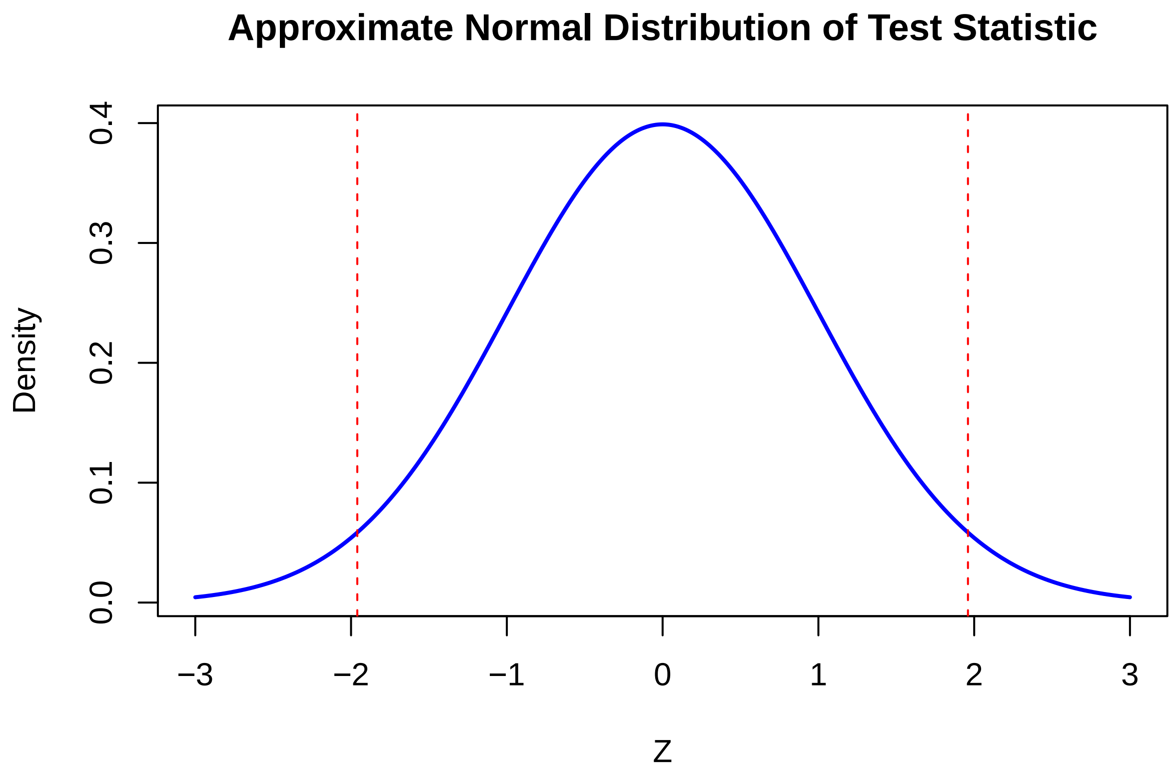
Alternative Hypothesis : H_1 , The coin is not fair & the probability of getting head (p) not equal to 0.5 ✓

i.e H_1 : $p \neq 0.5$

```r
# Part (b): Exact Distribution
n <- 20
p <- 0.5
x <- 0:n  # Possible number of heads
pmf <- dbinom(x, size = n, prob = p)  # Probability mass function

# Plotting the exact distribution
plot(x, pmf, type = "h", col = "blue", lwd = 2, xlab = "Number of Heads", ylab = "Probability", main =
abline(v = c(qbinom(0.025, size = n, prob = p), qbinom(0.975, size = n, prob = p)), col = "red", lty =
```




```r
# Part (c): Approximate Normal Distribution
sample_proportion <- 4 / n  # Number of heads observed
Z <- (sample_proportion - p) / sqrt(p * (1 - p) / n)

# Calculating p-value
p_value_clt <- 2 * pnorm(-abs(Z))

# Plotting the normal distribution
x_seq <- seq(-3, 3, length.out = 1000)
pdf_normal <- dnorm(x_seq, mean = 0, sd = 1)
plot(x_seq, pdf_normal, type = "l", col = "blue", lwd = 2, xlab = "Z", ylab = "Density", main = "Approxi
abline(v = c(-1.96, 1.96), col = "red", lty = 2)  # Critical region
```



✓

```r
# Part (d): Perform Test and Calculate p-values
observed_heads <- 4

# Binomial test (exact distribution)
p_value_binomial <- 2 * pbinom(observed_heads, size = n, prob = p, lower.tail = FALSE)

# CLT-based test
Z_observed <- (observed_heads / n - p) / sqrt(p * (1 - p) / n)
p_value_clt_observed <- 2 * pnorm(-abs(Z_observed))


cat("Exact Binomial Test p-value:", p_value_binomial, "\n")
```

this must be TRUE
then you will get the correct answer
note that

```
## Exact Binomial Test p-value: 1.988182
```

✗

p-value must be less than 1

```r
cat("CLT-based Test p-value:", p_value_clt_observed, "\n")
```

```
## CLT-based Test p-value: 0.007290358
```

✓

Q. 2 a

[0.9/1.0]

lets ~~Consider~~ denote the average price of shopping cart in the current year as μ. The null hypothesis (H_0) & alternative hypothesis (H_1) can be formulated as follows:-

$H_0: \mu \overset{<=}{=} 600$ (There is no change in the average price)

these two regions must cover all parameter space

$H_1: \mu > 600$ (The average price has increased)

```r
#FOS_HW13_Q2-b
# Part (b): Hypothesis Test
sample_mean <- 605
population_mean <- 600
standard_deviation <- 15
sample_size <- 40

# Calculating the test statistic (t-value)
t_value <- (sample_mean - population_mean) / (standard_deviation / sqrt(sample_size))

# Calculating the p-value
p_value <- pt(t_value, df = sample_size - 1, lower.tail = FALSE)

# Checking if we should reject the null hypothesis
alpha <- 0.05
reject_null <- p_value < alpha


cat("Test Statistic (t-value):", t_value, "\n")
```

```
## Test Statistic (t-value): 2.108185
```

```r
cat("p-value:", p_value, "\n")
```

```
## p-value: 0.020746
```

```r
cat("Reject H0 at level 0.05:", reject_null, "\n")
```

```
## Reject H0 at level 0.05: TRUE
```

```r
#FOS_HW13_Q2-c
# Part (c): Sample Size Calculation
desired_increase <- 5
desired_significance <- 0.01

# Z-score for 0.01 significance level (one-sided)
Z_alpha <- qnorm(1 - desired_significance)

# Calculating required sample size
required_sample_size <- ceiling((Z_alpha * standard_deviation / desired_increase)^2)


cat("Required Sample Size:", required_sample_size, "\n")
```

```
## Required Sample Size: 49
```

[0.8/0.8]

```r
#FOS_HW13_Q6
# Load the data
data(normtemp)
```

```
## Warning in data(normtemp): data set 'normtemp' not found
```

```r
# Viewing the structure of the data
str(normtemp)
```

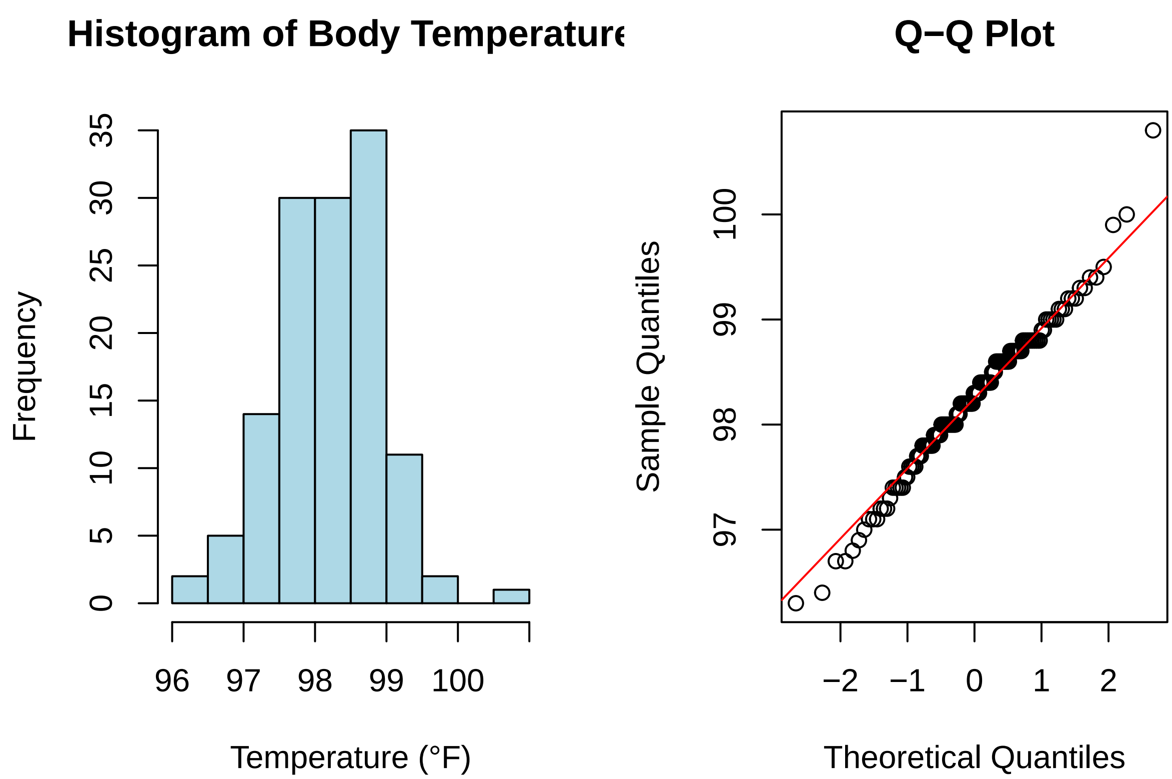
```
## 'data.frame':    130 obs. of  3 variables:
##  $ temperature: num  96.3 96.7 96.9 97 97.1 97.1 97.1 97.2 97.3 97.4 ...
##  $ gender     : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ hr         : int  70 71 74 80 73 75 82 64 69 70 ...
```

```r
# Summary statistics
summary(normtemp$temperature)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   96.30   97.80   98.30   98.25   98.70  100.80
```

✓

```r
# Checking the histogram and Q-Q plot
par(mfrow = c(1, 2))
hist(normtemp$temperature, main = "Histogram of Body Temperature", xlab = "Temperature (°F)", col = "li
qqnorm(normtemp$temperature, main = "Q-Q Plot")
qqline(normtemp$temperature, col = "red")
```



```r
par(mfrow = c(1, 1))



# One-sample t-test against the assumed value of 98.6°F
t_test_result <- t.test(normtemp$temperature, mu = 98.6)
t_test_result
```

```
##
##  One Sample t-test
##
## data:  normtemp$temperature
## t = -5.4548, df = 129, p-value = 2.411e-07
## alternative hypothesis: true mean is not equal to 98.6
## 95 percent confidence interval:
##  98.12200 98.37646
## sample estimates:
## mean of x
##  98.24923
```

No:

Date: __/__/__

05) i) $P(\text{I error}) = P(\text{Reject } H_0 \mid H_0 \text{ is true})$

[0.9/1.0] $P(\text{I error}) = P(|t| \geq 2 \mid \mu = 0)$

$P(\text{I error}) = P(t \leq -2 \text{ or } t \geq 2 \mid \mu = 0)$

using the symmetry of the standard normal distributn

$P(\text{I error}) = 2 \times P(t \geq 2 \mid \mu = 0)$

$P(t \geq 2 \mid \mu = 0)$
$= 0.0228$

$P(\text{I error}) = 2 \times 0.0228 = 0.0456$ ✓
4.56%

b) $H_0 = \mu = 0$ $H_1 = \mu \neq 0$

$\beta = P(\text{Fail to reject } H_0 \mid H_1 \text{ is true})$
$\beta = P(|t| < 2 \mid \mu = 1)$

$\beta = P(-2 < t < 2 \mid \mu = 1)$

$Z \sim N(\mu, 1)$;

$\beta = P\left(\frac{-2-1}{1} < Z < \frac{2-1}{1}\right)$
$\beta = P(-3 < Z < 1)$

$= 0.8413$ ✓ this is prob of type II error

Therefore the probability of committing a Type II error (β) when true value of μ is 1
~~approximately $1 - 0.8413 = 0.1587$ $\approx$ or 15.87%~~

[0.4/1.0] 3 (1) To determine if the rectangle's height to width ratios align with the golden ratio $\left(\frac{\sqrt{5}-1}{2}\right)$

first we can calculate the mean of the ratios and compare it with the golden ratio.

(1) Mean of the given ratios:

$$\text{Mean} = \frac{\text{Sum of ratios}}{\text{number of ratios}}$$

$$= \frac{0.693 + 0.749 + \ldots + 0.933}{20}$$

$$= \frac{13.27}{20} \cong 0.6635 \simeq 0.7$$

(2) Golden ratio $= \frac{\sqrt{5}-1}{2} \cong 0.61803 \simeq 0.6$

Comparing the mean with golden ratio we can see that they are close but not exactly the same. However, it's worth noting that the mean of the given ratio is slightly larger than the golden ratio.

$\therefore$ Based solely on this data, it's not entirely reasonable to conclude that the ancient peoples strictly followed the golden ratio in their rectangle designs. They may have been influenced by it, but further analysis would be needed to determine the extend of that influence yes one needs to do hypothesis testing!

(2) Bootstrap sample means :- we have 10 000 bootstrap sample means.

Mean and standard deviation of the original data $\bar{x}$:

$$\text{Mean} = \frac{13.27}{20} \simeq 0.6635$$

$$\text{Standard deviation} = \sum_{i=1}^{20} \sqrt{\frac{(x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{0.0646}{19}} \Rightarrow \sqrt{0.0034}$$

$$\cong 0.\cancel{0}5831$$

Standardized mean (t^*) for each bootstrap sample:

$$t^* = \frac{\bar{x}^*_n - \bar{x}_n}{s^*_n / \sqrt{n}}$$

$$t^* = \frac{\text{Bootstrap sample mean} - 0.6635}{0.\cancel{0}5891 / \sqrt{20}}$$

here you need to put bootstrap sample sd

After observing t^* values for all bootstrap samples, The 5^{th} percentile corresponds to the 500th value and the 95^{th} percentile corresponds to the 9500th value.

We count the number of t^* value that are greater than the absolute value of the standardize mean of the original data,

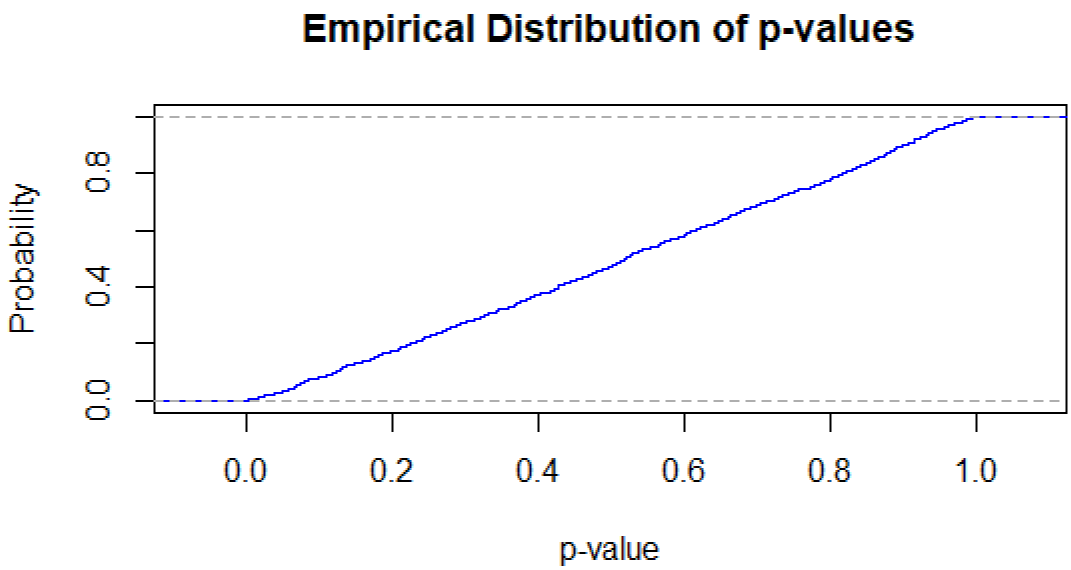
This count divided by the total number of t^* values gives us to the p value.

R code?

[1/1]

4)1

```
# Set the seed for reproducibility
> set.seed(123)
>
> # Draw a sample of size 100 from the standard normal distribution
> sample_data <- rnorm(100)
>
> # Perform two-sided t-test
> test_result <- t.test(sample_data, mu = 0)
>
> # Extract the p-value
> p_value <- test_result$p.value
>
> # Print the test result
> print(test_result)

        One Sample t-test

data:  sample_data
t = 0.99041, df = 99, p-value = 0.3244
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.09071657  0.27152838
sample estimates:
 mean of x
0.09040591

>
> # Print the p-value
> print(p_value)
[1] 0.3243898
> # Set the seed for reproducibility
> set.seed(123)
>
> # Initialize a vector to store p-values
> p_values <- numeric(1000)
>
> # Repeat the process a thousand times
> for (i in 1:1000) {
+     # Draw a sample of size 100 from the standard normal distribution
+     sample_data <- rnorm(100)
+
+     # Perform two-sided t-test and store the p-value
+     p_values[i] <- t.test(sample_data, mu = 0)$p.value
+ }
>
> # Plot the empirical distribution function of p-values
> plot(ecdf(p_values), main = "Empirical Distribution of p-values", xlab = "p
-value", ylab = "Probability", col = "blue")
```


Empirical Distribution of p-values
Probability
0.0
0.4
0.8
0.0
0.2
0.4
0.6
0.8
1.0
p-value