STAT 400 - Final Exam

Colin Gibbons-Fly

Question 1

```
# Given probabilities
P_A1 <- 0.30
P_A2 <- 0.20
P_A3 <- 0.50
P_B_given_A1 <- 0.03
P_B_given_A2 <- 0.05
P_B_given_A3 <- 0.04

# Calculate P(B) using the law of total probability
P_B <- (P_B_given_A1 * P_A1) + (P_B_given_A2 * P_A2) + (P_B_given_A3 * P_A3)

# Calculate P(A2 | B) using Bayes' Theorem
P_A2_given_B <- (P_B_given_A2 * P_A2) / P_B

# Display results
cat("P(A2 | B):", P_A2_given_B, "\n")
```

```
P(A2 | B): 0.2564103
```

Question 2

```
# A - Total executives and executives with >10 years
total_executives <- 200
exec_B4 <- 105

# Calculate P(B4)
```

```r
P_B4 <- exec_B4 / total_executives
cat("P(B4):", P_B4, "\n")
```

```
P(B4): 0.525
```

```r
# B - Executives not loyal and >10 years
exec_A2_B4 <- 30

# Calculate P(A2 | B4)
P_A2_given_B4 <- exec_A2_B4 / exec_B4
cat("P(A2 | B4):", P_A2_given_B4, "\n")
```

```
P(A2 | B4): 0.2857143
```

```r
# C -Total not loyal
exec_A2 <- 80

# Calculate P(A2)
P_A2 <- exec_A2 / total_executives

# Calculate P(B4   A2)
P_B4_union_A2 <- P_B4 + P_A2 - (exec_A2_B4 / total_executives)

#Display results
cat("P(B4   A2):", P_B4_union_A2, "\n")
```

```
P(B4   A2): 0.775
```

Question 3

Set-Up:

1) Null Hypothesis: True mean difference equals 0.
2) Alternative Hypothesis: True mean difference does not equal 0.

Decision Making Rules:

1) Significance Level = 0.05
2) Confidence Level = 95%

```r
# Data for Karlsruhe and Lehigh methods
karlsruhe <- c(1.186, 1.151, 1.322, 1.339, 1.200, 1.402, 1.365, 1.537, 1.559)
lehigh <- c(1.061, 0.992, 1.063, 1.062, 1.065, 1.178, 1.037, 1.086, 1.052)

# Calculate differences
differences <- karlsruhe - lehigh

# Summary statistics
mean_diff <- mean(differences)
sd_diff <- sd(differences)
n <- length(differences)

# Calculate the t statistic
t_stat <- mean_diff / (sd_diff / sqrt(n))

# Perform the paired t-test
t_test_result <- t.test(karlsruhe, lehigh, paired = TRUE)

# Display results
cat("Mean Difference:", mean_diff, "\n")
```

```
Mean Difference: 0.2738889
```

```r
cat("Standard Deviation of Differences:", sd_diff, "\n")
```

```
Standard Deviation of Differences: 0.1350994
```

```r
cat("t-Statistic:", t_stat, "\n")
```

```
t-Statistic: 6.081939
```

```r
print(t_test_result)
```

```
    Paired t-test

data:  karlsruhe and lehigh
t = 6.0819, df = 8, p-value = 0.0002953
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 0.1700423 0.3777355
sample estimates:
mean difference
      0.2738889
```

Decision:

1) Reject the Null Hypothesis. |t| > t_critical and the p-value is far below our significance level (0.05).

Conclusion:

At the 0.05 significance level we conclude there is a statistically significant difference between the means results of the Karlsruhe and Lehigh methods.

Question 4

```r
# Define the joint density function
joint_density <- function(x, y) {
  ifelse(x + y <= 1 & x >= 0 & y >= 0 & x <= 1 & y <= 1,
         24 * x * y,
         0)
}

# Define the marginal density of X
marginal_density_X <- function(x) {
  integrate(function(y) joint_density(x, y), 0, 1 - x)$value
}

# Define the conditional density of Y given X
conditional_density <- function(y, x) {
  joint_density(x, y) / marginal_density_X(x)
```

```r
}

# Compute the probability P(Y < 1/8 | X = 3/4)
probability <- integrate(function(y) conditional_density(y, 3/4), 0, 1/8)$value

# Display result
cat("The probability is:", probability, "\n")
```

```
The probability is: 0.25
```

Question 5

```r
# Define the joint probability distribution
prob_func <- function(x, y) {
  return((1 / (4^(x + y))) * (9 / 16))
}

# Compute E(X) using the definition of expectation
E_X <- sum(sapply(0:10, function(x) sum(sapply(0:10, function(y) x * prob_func(x, y)))))
# E(X) = E(Y) as X and Y are symmetric

# Compute E(Y) using the definition of expectation
E_Y <- sum(sapply(0:10, function(y) sum(sapply(0:10, function(x) y * prob_func(x, y)))))

# Compute the expected value of Z = X + Y
E_Z <- E_X + E_Y

# Compute E(X^2) for variance calculation
E_X2 <- sum(sapply(0:10, function(x) sum(sapply(0:10, function(y) (x^2) * prob_func(x, y))))
# Compute E(Y^2) in the same way
E_Y2 <- sum(sapply(0:10, function(y) sum(sapply(0:10, function(x) (y^2) * prob_func(x, y))))

# Compute the variance of X and Y
Var_X <- E_X2 - E_X^2
Var_Y <- E_Y2 - E_Y^2

# Since Var(Z) = Var(X) + Var(Y) (because X and Y are independent)
Var_Z <- Var_X + Var_Y
```

```
# Display results
cat("E(Z):", E_Z, "\n")
```

```
E(Z): 0.6666611
```

```
cat("Var(Z):", Var_Z, "\n")
```

```
Var(Z): 0.8888309
```

Question 6

```
# Define the total number of doctors and nurses
total_doctors <- 4
total_nurses <- 2
committee_size <- 3

# Define a function to calculate combinations (n choose k)
comb <- function(n, k) {
  return(choose(n, k))
}

# Calculate the total number of ways to select a committee
total_ways <- comb(total_doctors + total_nurses, committee_size)

# Calculate the probabilities for each possible number of doctors (x) on the committee
P_x_0 <- comb(total_doctors, 0) * comb(total_nurses, 3) / total_ways  # 0 doctors, 3 nurses
P_x_1 <- comb(total_doctors, 1) * comb(total_nurses, 2) / total_ways  # 1 doctor, 2 nurses
P_x_2 <- comb(total_doctors, 2) * comb(total_nurses, 1) / total_ways  # 2 doctors, 1 nurse
P_x_3 <- comb(total_doctors, 3) * comb(total_nurses, 0) / total_ways  # 3 doctors, 0 nurses

# Print the probabilities
cat("P(x = 0):", P_x_0, "\n")
```

```
P(x = 0): 0
```

```
cat("P(x = 1):", P_x_1, "\n")
```

```
P(x = 1): 0.2
```

```r
cat("P(x = 2):", P_x_2, "\n")
```

```
P(x = 2): 0.6
```

```r
cat("P(x = 3):", P_x_3, "\n")
```

```
P(x = 3): 0.2
```

```r
# Calculate P(2, 3)
P_2_3 <- P_x_2 + P_x_3

#Display result
cat("P(2 <= x <= 3):", P_2_3, "\n")
```

```
P(2 <= x <= 3): 0.8
```

Question 7

```r
# Given values
mean_X <- 18
sd_X <- 2.5
probability <- 0.8186  # 1 - 0.1814

# Find k using the inverse CDF (qnorm)
k <- qnorm(probability, mean = mean_X, sd = sd_X)

# Print the result
cat("The value of k such that P(X > k) = 0.1814 is:", k, "\n")
```

```
The value of k such that P(X > k) = 0.1814 is: 20.27511
```

```r
# Calculate P(X < 21) and P(X < 17)
p_21 <- pnorm(21, mean = mean_X, sd = sd_X)
p_17 <- pnorm(17, mean = mean_X, sd = sd_X)

# Calculate P(17 < X < 21)
p_17_21 <- p_21 - p_17
```

```
# Display result
cat("P(17 < X < 21) =", p_17_21, "\n")
```

```
P(17 < X < 21) = 0.5403521
```

Question 8

```
# Given values
alpha <- 1
beta <- 3

# Compute P(X > 1/3)
p_greater <- pbeta(1/3, alpha, beta)
cat("P(X > 1/3):", 1 - p_greater, "\n")
```

```
P(X > 1/3): 0.2962963
```

Question 9

```
# Given values
mu_xbar <- 50  # population mean
sigma_xbar <- 5 / sqrt(16)  # standard error of the sample mean
lower_bound <- mu_xbar - 1.96 * sigma_xbar
upper_bound <- mu_xbar - 0.4 * sigma_xbar

# Standardize the bounds to Z-scores
Z_lower <- (lower_bound - mu_xbar) / sigma_xbar
Z_upper <- (upper_bound - mu_xbar) / sigma_xbar

# Calculate the probabilities using the cumulative distribution function (pnorm)
P_lower <- pnorm(Z_lower)
P_upper <- pnorm(Z_upper)

# The probability that x_bar falls between the two bounds
P_interval <- P_upper - P_lower

# Display result
cat("The probability that the sample mean falls between the bounds is:", P_interval, "\n")
```

```
The probability that the sample mean falls between the bounds is: 0.3195804
```

Question 10

```r
# Given values
mu <- 7950  # Population mean (government limit)
X_bar <- 7960  # Sample mean
s <- 100  # Population standard deviation
n <- 25  # Sample size

# Standard error of the mean
sigma_Xbar <- s / sqrt(n)

# Z-score calculation
z <- (X_bar - mu) / sigma_Xbar

# Probability that X_bar >= 7960 assuming mu = 7950
probability <- 1 - pnorm(z)

# Display result
cat("P(X >= 7960 | mu = 7950):", probability, "\n")
```

```
P(X >= 7960 | mu = 7950): 0.3085375
```

Question 11

```r
# Given values
s <- 3900  # Sample standard deviation
n <- 100   # Sample size
confidence_level <- 0.99  # Confidence level

# Critical value for 99% confidence level
z_critical <- qnorm(1 - (1 - confidence_level) / 2)

# Standard error of the mean
standard_error <- s / sqrt(n)

# Margin of error
me <- z_critical * standard_error
```

```r
# Display result
cat("Margin of Error (99% confidence):", me, "\n")
```

```
Margin of Error (99% confidence): 1004.573
```

Conclusion

With 99% confidence, we can assert that the true mean kilometers driven by car owners in Maryland lies within 22,495.427 kilometers and 24,504.573 kilometers. The possible size of the error in this estimate is ±1004.573 kilometers.

Question 12

```r
# Given data
x1_bar <- 36300  # Brand A mean
s1 <- 5000       # Brand A standard deviation
n1 <- 12         # Brand A sample size

x2_bar <- 38100  # Brand B mean
s2 <- 6100       # Brand B standard deviation
n2 <- 12         # Brand B sample size

# Standard Error calculation
se <- sqrt((s1^2 / n1) + (s2^2 / n2))

# Degrees of freedom calculation (Welch-Satterthwaite equation)
df <- ((s1^2 / n1 + s2^2 / n2)^2) / (((s1^2 / n1)^2 / (n1 - 1)) + ((s2^2 / n2)^2 / (n2 - 1))

# t-critical value for 95% confidence level
t_critical <- qt(1 - 0.05 / 2, df)

# Margin of error
me <- t_critical * se

# Confidence interval
CI_lower <- (x1_bar - x2_bar) - me
CI_upper <- (x1_bar - x2_bar) + me
```

```
# Display result
cat("95% Confidence Interval: (", CI_lower, ",", CI_upper, ")\n")
```

```
95% Confidence Interval: ( -6532.522 , 2932.522 )
```

Question 13

```
# Given values
p_hat <- 2 / 3  # Sample proportion
n <- 1600       # Sample size
confidence_level <- 0.95  # Confidence level

# Standard error calculation
se <- sqrt(p_hat * (1 - p_hat) / n)

# Z-critical value for 95% confidence level
z_critical <- qnorm(1 - (1 - confidence_level) / 2)

# Margin of error
me <- z_critical * se

# Confidence interval
CI_lower <- p_hat - me
CI_upper <- p_hat + me

# Display results
cat("95% Confidence Interval: (", CI_lower, ",", CI_upper, ")\n")
```

```
95% Confidence Interval: ( 0.6435683 , 0.6897651 )
```

```
cat("Margin of Error:", me, "\n")
```

```
Margin of Error: 0.0230984
```

Conclusion

With 95% confidence, the true proportion of American adults who think the space shuttle program is a good investment is likely to be within ± 0.03 of the sample estimate (2/3 or

0.6667). In other words, the possible size of the error in estimating the population proportion is about 0.03.

If we were to repeatedly sample from the population, 95% of the time, the true population proportion would lie within the range of 0.6436 to 0.6898. This range reflects the uncertainty in our estimate and provides a margin of error of approximately 3%.

Question 14

```r
# Given values
sigma1_sq <- 0.6667 * 0.3333 / 1600 # Variance of group 1
sigma2_sq <- 0.6667 * 0.3333 / 1600 # Variance of group 2
n1 <- 1600
n2 <- 1600
alpha <- 0.05 # For a 95% confidence interval

# Degrees of freedom
df1 <- n1 - 1
df2 <- n2 - 1

# Variance ratio
var_ratio <- sigma1_sq / sigma2_sq

# Critical values for F-distribution
f_upper <- qf(1 - alpha / 2, df1, df2)
f_lower <- qf(alpha / 2, df1, df2)

# Confidence interval
lower <- var_ratio / f_upper
upper <- var_ratio / f_lower

#Display results
cat("Variance Ratio:", var_ratio, "\n")
```

```
Variance Ratio: 1
```

```r
cat("95% Confidence Interval for Variance Ratio: (", lower, ",", upper, ")\n")
```

```
95% Confidence Interval for Variance Ratio: ( 0.9065911 , 1.103033 )
```

Conclusion

Given that the confidence interval includes 1 and the variance ratio is exactly 1, there is no statistical evidence to suggest $\sigma_1^2 \neq \sigma_2^2$. Therefore, the assumption of unequal variances when constructing the confidence intervals was not justified.

Question 15

```
# Given values
mu_null <- 5000          # Null hypothesis mean
mu_4970 <- 4970          # Alternative hypothesis mean (mu = 4970)
mu_4960 <- 4960          # Alternative hypothesis mean (mu = 4960)
s <- 120                 # Standard deviation
n <- 50                  # Sample size
critical_value <- 4790   # Critical value for the sample mean

# Standard error of the mean
se <- s / sqrt(n)

# Z-scores for the critical value under both alternatives
z_4970 <- (critical_value - mu_4970) / se
z_4960 <- (critical_value - mu_4960) / se

# Calculate the probability of Type II error (beta)
beta_4970 <- pnorm(z_4970)  # P(Z >= z_4970)
beta_4960 <- pnorm(z_4960)  # P(Z >= z_4960)

# Print results
cat("For mu = 4970, Z-score =", z_4970, "and beta =", beta_4970, "\n")
```

```
For mu = 4970, Z-score = -10.6066 and beta = 1.388325e-26
```

```
cat("For mu = 4960, Z-score =", z_4960, "and beta =", beta_4960, "\n")
```

```
For mu = 4960, Z-score = -10.01735 and beta = 6.394549e-24
```

Question 16

```r
# Given values
x_bar <- 23500      # Sample mean
mu_0 <- 20000       # Population mean under the null hypothesis
s <- 3900       # Sample standard deviation
n <- 100            # Sample size

# Standard error of the mean
se <- s / sqrt(n)

# Z-score
z <- (x_bar - mu_0) / se

# P-value (one-tailed)
p_value <- pt(t_stat, df = n - 1, lower.tail = FALSE)

# Display results
cat("Z-score:", z, "\n")
```

```
Z-score: 8.974359
```

```r
cat("P-Value:", p_value, "\n")
```

```
P-Value: 1.119986e-08
```

Conclusion

Since the p-value is significantly less than 0.05, we reject the null hypothesis and conclude that there is strong evidence to support the claim that automobiles are driven on average more than 20,000 kilometers per year.

Question 17

```r
# Given data
treatment <- c(2.1, 5.3, 1.4, 4.6, 0.9)
no_treatment <- c(1.9, 0.5, 2.8, 3.1)

# Perform t-test
t_test <- t.test(treatment, no_treatment, alternative = "greater", var.equal = TRUE)
```

```r
# Display results
print(t_test)
```

```
    Two Sample t-test

data:  treatment and no_treatment
t = 0.69899, df = 7, p-value = 0.2536
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.342717       Inf
sample estimates:
mean of x mean of y
    2.860     2.075
```

Conclusion

Given the p-value is greater than 0.05, we fail to reject the null hypothesis and conclude that there is insufficient evidence to say the serum is effective at arresting leukemia.

Question 18

```r
# Given data
s1_sq <- 6.1^2 # Variance for men
s2_sq <- 5.3^2 # Variance for women
n1 <- 11
n2 <- 14

# Test statistic
F <- s1_sq / s2_sq

# Degrees of freedom
df1 <- n1 - 1
df2 <- n2 - 1

# P-value for one-tailed test
p_value <- pf(F, df1, df2, lower.tail = FALSE)
```

```r
#Display results
cat("Test Statistic (F):", F, "\n")
```

```
Test Statistic (F): 1.324671
```

```r
cat("P-Value:", p_value, "\n")
```

```
P-Value: 0.3118285
```

Conclusion

Given the p-value is greater than 0.05, we fail to reject the null hypothesis and conclude that there is insufficient evidence to say the variance of assembly times for men is significantly greater than that for women.

Question 19

```r
# Given data: Number of advertisements (X) and Total Sales (Y)
X <- c(20, 40, 20, 30, 10, 10, 20, 20, 20, 30)
Y <- c(30, 60, 40, 60, 30, 40, 40, 50, 30, 70)

# Number of data points
n <- length(X)

# (a) Calculate the correlation coefficient (r)
r <- cor(X, Y)

# (b) Test the null hypothesis that   = -0.5 against the alternative   < -0.5
# Compute the t-statistic for correlation test
t_statistic <- (r * sqrt(n - 2)) / sqrt(1 - r^2)

# Critical value for t-distribution with (n-2) degrees of freedom, one-tailed, alpha = 0.025
critical_value <- qt(0.025, df = n - 2)

# (c) Find the equation of the regression line
lm_model <- lm(Y ~ X)
b0 <- coef(lm_model)[1]
b1 <- coef(lm_model)[2]
```

```r
# (d) Estimate the amount of revenue generated when 22 advertisements are made
X_new <- 22
Y_pred <- b0 + b1 * X_new

# (e) Determine the percentage of the variation (coefficient of determination)
r_squared <- summary(lm_model)$r.squared * 100  # Convert to percentage

# (f) Find the standard error of the estimates
std_error <- summary(lm_model)$sigma

# (g) Test the hypothesis of the regression coefficient at 0.01 level of significance
# Null hypothesis: H0:  1 = 0 (no relationship)
# Alternative hypothesis: H1:  1   0 (there is a relationship)
p_value_b1 <- summary(lm_model)$coefficients[2, 4]  # p-value for b1

# (h) Find the F-value
f_value <- summary(lm_model)$fstatistic[1]

# Display results
cat("Correlation coefficient (r):", r, "\n")
```

```
Correlation coefficient (r): 0.7590141
```

```r
cat("t-statistic:", t_statistic, "\n")
```

```
t-statistic: 3.297345
```

```r
cat("Critical value:", critical_value, "\n")
```

```
Critical value: -2.306004
```

```r
cat("Regression equation: Y =", b0, "+", b1, "* X\n")
```

```
Regression equation: Y = 18.94737 + 1.184211 * X
```

```r
cat("Estimated revenue for 22 advertisements:", Y_pred, "\n")
```

```
Estimated revenue for 22 advertisements: 45
```

```r
cat("Percentage of variation (R-squared):", r_squared, "%\n")
```

```
Percentage of variation (R-squared): 57.61024 %
```

```r
cat("Standard error of the estimate:", std_error, "\n")
```

```
Standard error of the estimate: 9.900824
```

```r
cat("p-value for regression coefficient (b1):", p_value_b1, "\n")
```

```
p-value for regression coefficient (b1): 0.01090193
```

```r
cat("F-value:", f_value, "\n")
```

```
F-value: 10.87248
```

Question 20

```r
# Given data
data <- data.frame(
  Heating_Cost = c(250, 360, 155, 43, 92, 200, 355, 290, 230, 120, 73, 205, 400, 320, 72, 27
  Mean_Outside_Temperature = c(35, 29, 36, 20, 65, 30, 10, 7, 21, 55, 54, 48, 20, 39, 60, 20
  Attic_Insulation = c(3, 4, 7, 6, 5, 5, 6, 10, 9, 2, 12, 5, 5, 4, 8, 5, 7, 8, 9, 7),
  Age_of_Furnace = c(6, 10, 3, 9, 6, 5, 7, 10, 11, 5, 4, 1, 15, 7, 6, 8, 3, 11, 8, 5)
)

# Fit the multiple linear regression model
model <- lm(Heating_Cost ~ Mean_Outside_Temperature + Attic_Insulation + Age_of_Furnace, data
print(summary(model))
```

```
Call:
lm(formula = Heating_Cost ~ Mean_Outside_Temperature + Attic_Insulation +
    Age_of_Furnace, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-232.72  -21.70    0.57   39.10   95.34

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)               347.567     94.243   3.688  0.00199 **
Mean_Outside_Temperature   -3.249      1.272  -2.555  0.02117 *
Attic_Insulation          -13.727      7.139  -1.923  0.07250 .
Age_of_Furnace              8.389      6.438   1.303  0.21099
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 76.77 on 16 degrees of freedom
Multiple R-squared:  0.559, Adjusted R-squared:  0.4763
F-statistic: 6.761 on 3 and 16 DF,  p-value: 0.00372
```

```r
# (a) Estimated Regression Equation
cat("Estimated Regression Equation:\nHeating_Cost =",
    round(coef(model)[1], 3), "+",
    round(coef(model)[2], 3), "* Mean_Outside_Temperature +",
    round(coef(model)[3], 3), "* Attic_Insulation +\n",
    round(coef(model)[4], 3), "* Age_of_Furnace\n")
```

```
Estimated Regression Equation:
Heating_Cost = 347.567 + -3.249 * Mean_Outside_Temperature + -13.727 * Attic_Insulation +
 8.389 * Age_of_Furnace
```

```r
# (b) Coefficients Interpretation
coefficients <- summary(model)$coefficients
print(coefficients)
```

```
                           Estimate Std. Error   t value    Pr(>|t|)
(Intercept)              347.567168  94.243482  3.687970 0.001992346
Mean_Outside_Temperature  -3.249255   1.271534 -2.555382 0.021171576
Attic_Insulation         -13.727054   7.139424 -1.922712 0.072502805
Age_of_Furnace             8.389057   6.437937  1.303066 0.210993948
```

```r
# (c) Heating cost estimate for specific conditions
predicted_cost <- predict(model, newdata = data.frame(Mean_Outside_Temperature = 35,
                                                      Attic_Insulation = 3,
                                                      Age_of_Furnace = 6))
cat("\nPredicted Heating Cost for T=35, Insulation=3, Age=6:", predicted_cost, "\n")
```

```
Predicted Heating Cost for T=35, Insulation=3, Age=6: 242.9964
```

```r
# (d) Multiple Standard Error of the Estimates
se <- summary(model)$sigma
cat("\nMultiple Standard Error of the Estimates:", se, "\n")
```

```
Multiple Standard Error of the Estimates: 76.76534
```

```r
# (e) Coefficient of Multiple Determination (R-squared)
r_squared <- summary(model)$r.squared
cat("\nCoefficient of Multiple Determination (R-squared):", r_squared, "\n")
```

```
Coefficient of Multiple Determination (R-squared): 0.5590283
```

```r
# (f) Adjusted Coefficient of Determination (Adjusted R-squared)
adjusted_r_squared <- summary(model)$adj.r.squared
cat("\nAdjusted Coefficient of Determination (Adjusted R-squared):", adjusted_r_squared, "\n
```

```
Adjusted Coefficient of Determination (Adjusted R-squared): 0.4763461
```

```r
# (g) Hypothesis Test for Regression Coefficients
cat("\nHypothesis Test for Regression Coefficients (P-values):\n")
```

```
Hypothesis Test for Regression Coefficients (P-values):
```

```r
p_values <- summary(model)$coefficients[, 4]
print(p_values)
```

```
            (Intercept) Mean_Outside_Temperature          Attic_Insulation
            0.001992346              0.021171576               0.072502805
         Age_of_Furnace
            0.210993948
```

```r
# (h) t-test for Temperature and Insulation Variables
p_value_temp <- summary(model)$coefficients[2, 4]
p_value_insul <- summary(model)$coefficients[3, 4]

cat("\nP-value for Mean Outside Temperature:", p_value_temp, "\n")
```

```
P-value for Mean Outside Temperature: 0.02117158
```

```r
cat("P-value for Attic Insulation:", p_value_insul, "\n")
```

```
P-value for Attic Insulation: 0.07250281
```

```r
significance_temp <- ifelse(p_value_temp < 0.05, "Significant", "Not Significant")
significance_insul <- ifelse(p_value_insul < 0.05, "Significant", "Not Significant")

cat("\nSignificance of Temperature Variable:", significance_temp, "\n")
```

```
Significance of Temperature Variable: Significant
```

```r
cat("Significance of Insulation Variable:", significance_insul, "\n")
```

```
Significance of Insulation Variable: Not Significant
```

```r
cat("Estimated Regression Equation: Heating_Cost = ",
    round(coef(model)[1], 3), " + ",
    round(coef(model)[2], 3), " * Mean_Outside_Temperature + ",
    round(coef(model)[3], 3), " * Attic_Insulation + ",
    round(coef(model)[4], 3), " * Age_of_Furnace\n")
```

```
Estimated Regression Equation: Heating_Cost =  347.567  +  -3.249  * Mean_Outside_Temperatur
```

I have completed this quiz myself, working independently and not consulting anyone except the instructor. I have neither given nor received help on this quiz. Name: Colin Gibbons-Fly Date: December 10, 2024