Full name: Nguyen Nhut Vu Truong

Student ID: 413707008

Course: Financial Econometrics

HW0310

Question 1

(a) The null hypothesis is $H_0 : \beta_2 = 0$ and the alternative hypothesis is $H_1 : \beta_2 > 0$.

(b) The test statistic is $t = b_2 / \text{se}(b_2)$. If the null hypothesis is true then $t \sim t_{(62)}$.

(c) Under the alternative hypothesis the center of the *t*-distribution is pushed to the right.

(d) We will reject the null hypothesis and accept the alternative if $t \geq 2.388$. We fail to reject the null hypothesis if $t < 2.388$.

(e) The calculated value of the test statistic is $t = 6.0884$. We reject the null hypothesis that there is no relationship between the number of medals won and *GDP* and we accept the alternative that there is positive relationship between the number of medals won and *GDP*. The level of significance of a test *is* the probability of committing a Type I error.

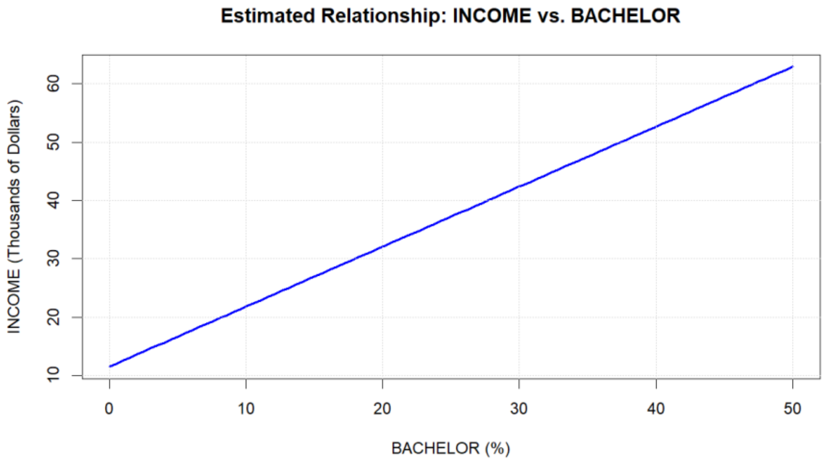
```
> # (b) Test statistic:
> t_stat <- b2 / se_b2
> cat("Test statistic (t):", t_stat, "\n")
Test statistic (t): 6.088372
>
> # Under H0, t ~ t_{df=62}
>
> # (c) Critical value at 1% (one-sided):
> alpha  <- 0.01
> t_crit <- qt(1 - alpha, df)
> cat("Critical value at 1% one-sided:", t_crit, "\n")
Critical value at 1% one-sided: 2.388011
>
> # (d) Decision:
> if(t_stat > t_crit) {
+   cat("Reject H0 at the 1% level.\n")
+ } else {
+   cat("Fail to reject H0 at the 1% level.\n")
+ }
Reject H0 at the 1% level.
>
> # p-value (one-sided):
> p_val <- pt(t_stat, df, lower.tail = FALSE)
> cat("One-sided p-value:", p_val, "\n")
One-sided p-value: 3.943571e-08
>
> # Economic conclusion:
> cat("Conclusion: There is a statistically significant positive relationship between GDP and medals.\n")
Conclusion: There is a statistically significant positive relationship between GDP and medals.
```

```
> # (e) The level of significance of a test is the probability of committing a Type I error.
> # Compute t-statistic
> t_stat <- (b2 - 0) / se_b2
> cat("Test statistic (t-value):", t_stat, "\n")
Test statistic (t-value): 402.1578
>
> # Decision on hypothesis test
> if (t_stat > t_crit) {
+   cat("Conclusion: Reject H0. GDP is significantly positively associated with medals won at the 1% level.\n")
+ } else {
+   cat("Conclusion: Fail to reject H0. There is not enough evidence to conclude a positive relationship at the 1% level.\n")
+ }
Conclusion: Reject H0. GDP is significantly positively associated with medals won at the 1% level.
```

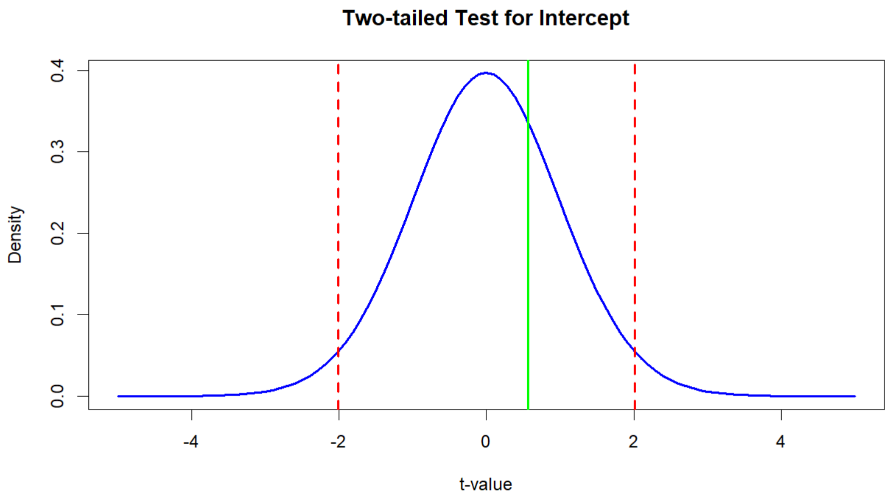
Question 7

```
> # (a) Calculate the estimated intercept
> b0 <- t_b0 * se_b0  # Intercept formula: b0 = t * SE
> cat("Estimated intercept (b0):", b0, "\n")
Estimated intercept (b0): 11.51632
>
> # (b) Sketch the estimated relationship
> bachelor_values <- seq(0, 50, length.out = 100)  # Generate BACHELOR values
> income_values <- b0 + b1 * bachelor_values  # Compute estimated INCOME values
>
> plot(bachelor_values, income_values, type="l", col="blue", lwd=2,
+      xlab="BACHELOR (%)", ylab="INCOME (Thousands of Dollars)",
+      main="Estimated Relationship: INCOME vs. BACHELOR")
> grid()
```

The relationship between INCOME and BACHELOR is increasing at a constant rate.




```
> # (c) Calculate the standard error of the slope coefficient
> # Formula: SE = b1 / t_b1
> se_b1 <- b1 / t_b1
> cat("Calculated standard error of slope:", se_b1, "\n")
Calculated standard error of slope: 0.09572093
>
> # (d) Calculate t-statistic for testing H0: intercept = 10
> # Formula: t = (b0 - 10) / se_b0
> t_stat_intercept <- (b0 - 10) / se_b0
> cat("t-statistic for testing intercept=10:", t_stat_intercept, "\n")
t-statistic for testing intercept=10: 0.567485
>
> # (e) Compute p-value for two-tailed test of intercept = 10
> p_value_intercept <- 2 * (1 - pt(abs(t_stat_intercept), df))
> cat("p-value for intercept test:", p_value_intercept, "\n")
p-value for intercept test: 0.5729757
>
> # Sketch the rejection region
> curve(dt(x, df), from=-5, to=5, col="blue", lwd=2,
+       xlab="t-value", ylab="Density",
+       main="Two-tailed Test for Intercept")
> qt(0.975, df = df) # Two-tailed: 0.05/2 = 0.025, t value = 2.0096
[1] 2.009575
> abline(v = c(-2.0096, 2.0096), col="red", lwd=2, lty=2)  # Critical values
> abline(v = t_stat_intercept, col="green", lwd=2)  # Computed t-value
```




```
> # (f) Construct a 99% confidence interval for the slope
> alpha_99 <- 0.01
> t_crit_99 <- qt(1 - alpha_99/2, df)
> CI_99_slope <- b1 + c(-1, 1) * t_crit_99 * se_b1
> cat("99% Confidence Interval for slope:", CI_99_slope, "\n")
99% Confidence Interval for slope: 0.7724725 1.285527
>
> # (g) Test H0: slope = 1 at 5% significance level
> t_stat_slope <- (b1 - 1) / se_b1
> p_value_slope <- 2 * (1 - pt(abs(t_stat_slope), df))
> cat("t-statistic for testing slope=1:", t_stat_slope, "\n")
t-statistic for testing slope=1: 0.302964
> cat("p-value for testing slope=1:", p_value_slope, "\n")
p-value for testing slope=1: 0.7631998
>
> # Decision on hypothesis test
> if (p_value_slope < 0.05) {
+   cat("Conclusion: Reject H0; slope is significantly different from 1.\n")
+ } else {
+   cat("Conclusion: Fail to reject H0; slope is not significantly different from 1.\n")
+ }
Conclusion: Fail to reject H0; slope is not significantly different from 1.
```

Question 17

```
> # a. Hypothesis testing for the urban regression
> # Null Hypothesis: H0: beta2 = 1.80
> # Alternative Hypothesis: H1: beta2 > 1.80
> alpha <- 0.05
> t_statistic <- (urban_beta2 - 1.80) / urban_se_beta2
> critical_value <- qt(1 - alpha, df = N_urban - 2)  # One-tailed test
>
> cat("a. t-statistic:", round(t_statistic, 3), "\n")
a. t-statistic: 4.125
> cat("Critical value at alpha =", alpha, ":", round(critical_value, 3), "\n")
Critical value at alpha = 0.05 : 1.646
> cat("Conclusion: ", ifelse(t_statistic > critical_value, "Reject H0: t falls in the rejection region, so we
+ reject the null hypothesis and accept the alternative.", "Fail to reject H0"), "\n\n")
Conclusion:  Reject H0: t falls in the rejection region, so we
reject the null hypothesis and accept the alternative.
```

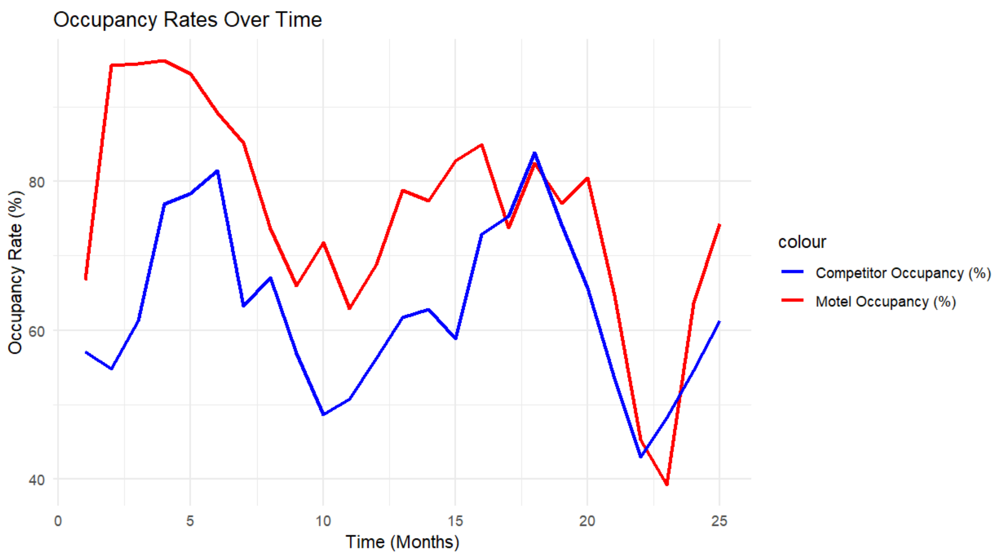
```
> # b. 95% confidence interval for rural regression
> expected_WAGE <- rural_beta1 + rural_beta2 * 16
> std_error <- 0.833  # Given standard error
> covariance <- -0.761  # Given covariance
> std_error_rural <- sqrt((rural_se_beta1^2) + 16^2*(rural_se_beta2^2) + (2 * 16 * covariance))
> std_error_rural
[1] 1.103494
> # Confidence interval calculation
> margin_of_error <- critical_value_b * std_error
> ci_lower <- expected_WAGE - margin_of_error
> ci_upper <- expected_WAGE + margin_of_error
> cat("b. 95% CI for expected WAGE if EDUC = 16: [", round(ci_lower, 2), ",", round(ci_upper, 2), "]\n\n")
b. 95% CI for expected WAGE if EDUC = 16: [ 22.28 , 25.56 ]

> # c. 95% confidence interval for expected WAGE using urban regression
> expected_WAGE_urban <- urban_beta1 + urban_beta2 * 16
> covariance <- -0.345
> std_error_c <- sqrt((urban_se_beta1^2) + 16^2*(urban_se_beta2^2) + (2 * 16 * covariance))  # Combined SE
> critical_value_c <- qt(0.975, df = N_urban - 2)
> # Confidence interval calculation
> margin_of_error_c <- critical_value_c * std_error_c
> ci_lower_c <- expected_WAGE_urban - margin_of_error_c
> ci_upper_c <- expected_WAGE_urban + margin_of_error_c
> cat("c. 95% CI for expected WAGE if EDUC = 16 (urban): [", round(ci_lower_c, 2), ",", round(ci_upper_c, 2), "]\n
\n")
c. 95% CI for expected WAGE if EDUC = 16 (urban): [ 27 , 30.2 ]

> # d. Hypothesis testing for the rural regression intercept
> # Null Hypothesis: H0: beta1 >= 4
> # Alternative Hypothesis: H1: beta1 < 4
> t_statistic_d <- (rural_beta1-4) / rural_se_beta1
> critical_value_d <- qt(alpha, df = N_rural - 2)  # One-tailed test
>
> cat("d. t-statistic for intercept:", round(t_statistic_d, 2), "\n")
d. t-statistic for intercept: -2.7
> cat("Critical value at alpha =", alpha, ":", round(critical_value_d, 3), "\n")
Critical value at alpha = 0.05 : -1.652
> cat("Conclusion: ", ifelse(t_statistic_d < critical_value_d, "Reject H0: t falls in the rejection region, so we
+ reject the null hypothesis and accept the alternative.", "Fail to reject H0"), "\n")
Conclusion:  Reject H0: t falls in the rejection region, so we
reject the null hypothesis and accept the alternative.
```

Question 19

a. Yes, they tend to move together. Motel seems to have the higher occupancy rates.



```
> # Estimate the regression model
> reg_model <- lm(motel_pct ~ comp_pct, data = motel)
> summary_reg <- summary(reg_model)
> summary(reg_model)

Call:
lm(formula = motel_pct ~ comp_pct, data = motel)

Residuals:
    Min      1Q  Median      3Q     Max 
-23.876  -4.909  -1.193   5.312  26.818 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  21.4000    12.9069   1.658 0.110889    
comp_pct      0.8646     0.2027   4.265 0.000291 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.02 on 23 degrees of freedom
Multiple R-squared:  0.4417,	Adjusted R-squared:  0.4174 
F-statistic: 18.19 on 1 and 23 DF,  p-value: 0.0002906
```

```
> # 95% confidence interval for the slope (beta_2)
> conf_int <- confint(reg_model, level = 0.95)
> cat("95% CI for beta_2:", round(conf_int[2, 1], 4), "to", round(conf_int[2, 2], 4), "\n")
95% CI for beta_2: 0.4453 to 1.284
```

$$\widehat{MOTEL\_PCT} = 21.340 + 0.865 COMP\_PCT$$
$$\text{(se)} \qquad (12.907) \quad (0.203)$$

A 95% interval estimate for β_2 is [0.445, 1.284].

The estimate of the association between MOTEL_PCT and COMP_PCT is positive and statistically significant, but it is not estimated with high precision due to the moderately wide confidence interval and non-trivial standard error. If the interval were narrower, we could be more confident about the precise effect of COMP_PCT on MOTEL_PCT.

```
> # b. Construct a 90% interval estimate for expected MOTEL_PCT given COMP_PCT = 70
> # Required parameters
> alpha <- 0.10  # for 90% CI
> tcr <- qt(1 - alpha / 2, df = df.residual(model_b))  # Critical value
> b1 <- coef(model_b)[1]  # Intercept
> b2 <- coef(model_b)[2]  # Slope
> varb1 <- vcov(model_b)[1, 1]  # Variance of intercept
> varb2 <- vcov(model_b)[2, 2]  # Variance of slope
> covb12 <- vcov(model_b)[1, 2]  # Covariance
> 
> # Calculate the expected value for COMP_PCT = 70
> comp_pct_value <- 70
> expected_value <- b1 + b2 * comp_pct_value
> 
> # Calculate the standard error for the estimated value
> varL <- varb1 + (comp_pct_value^2 * varb2) + (2 * comp_pct_value * covb12)  # Variance of the linear estimate
> seL <- sqrt(varL)  # Standard error of L
> 
> # Calculate the confidence interval
> lowbL <- expected_value - tcr * seL
> upbL <- expected_value + tcr * seL
> 
> cat("b. 90% CI for expected MOTEL_PCT given COMP_PCT =", comp_pct_value, ": [",
+     round(lowbL, 3), ",", round(upbL, 3), "]\n\n")
b. 90% CI for expected MOTEL_PCT given COMP_PCT = 70 : [ 77.382 , 86.467 ]
```

```
> # c. Test the null hypothesis H0: β2 ≤ 0 against H1: β2 > 0 at the α = 0.01 level of significance
> t_statistic_c <- (b2 - 0) / sqrt(varb2)  # Calculate t-statistic
> critical_value_c <- qt(1 - 0.01, df = df.residual(model_b))  # One-tailed test
>
> cat("c. t-statistic for β2:", round(t_statistic_c, 2), "\n")
c. t-statistic for β2: 4.27
> cat("Critical value at α = 0.01:", round(critical_value_c, 3), "\n")
Critical value at α = 0.01: 2.5
> cat("Conclusion: ", ifelse(t_statistic_c > critical_value_c, "Reject H0", "Fail to reject H0"), "\n\n")
Conclusion:  Reject H0
```

```
> # d. Test the null hypothesis H0: β2 = 1 against H1: β2 ≠ 1 at the α = 0.01 level of significance
> t_statistic_d <- (b2 - 1) / sqrt(varb2)  # Calculate t-statistic for β2 = 1
> critical_value_d <- qt(1 - 0.005, df = df.residual(model_b))  # Two-tailed test
>
> cat("d. t-statistic for β2 = 1:", round(t_statistic_d, 2), "\n")
d. t-statistic for β2 = 1: -0.67
> cat("Critical value at α = 0.01 (two-tailed):", round(critical_value_d, 3), "\n")
Critical value at α = 0.01 (two-tailed): 2.807
> cat("Conclusion: ", ifelse(abs(t_statistic_d) > critical_value_d, "Reject H0", "Fail to reject H0"), "\n")
Conclusion:  Fail to reject H0
```

```
> residuals_e
          1           2           3           4           5           6           7           8           9 
 -4.0708929  26.8177776  21.3976220   8.3092488   5.3122899  -2.5816281   9.1548074  -5.7172859  -4.5979650 
         10          11          12          13          14          15          16          17          18 
  8.3785413  -2.3372013  -1.1927175   4.0517663   1.7006631  10.4727564   0.4678061 -12.7073283 -11.5432263 
         19          20          21          22          23          24          25 
 -8.4562250   2.2796730  -2.9581913 -13.2930147 -23.8756030  -4.9092946  -0.1023780 
```

The residual plot indeed shows high variability both in the early time periods and toward the end. This suggests potential instability in the relationship between MOTEL_PCT and COMP_PCT over time.

For observations 17-23 all the residuals are negative but one.