Homework 1

2023-09-11

```r
#A

library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
# Loading the data
data <- read_excel("Houses.xlsx")
# Filtering non-traditional-style houses
non_traditional_data <- data %>% filter(Traditional == 0)

# Linear regression
model_1 <- lm(price ~ sqft, data = non_traditional_data)

# Performing the hypothesis test

# Getting the p-value associated with the sqft variable (slope coefficient)
p <- summary(model_1)$coefficients["sqft", "Pr(>|t|)"]

# Setting the significance level (alpha)
a <- 0.05
# Checking if p-value is less than alpha
if (p < a) {
  print("Reject the null hypothesis")
} else {
  print("Fail to reject the null hypothesis")
}
```

```
## [1] "Reject the null hypothesis"
```

```r
#B

# Predicting the price for a house with 1500 square feet
predicted_price <- predict(model_1, newdata = data.frame(sqft = 1500))
# Setting the null hypothesis value
expected_price <- 90000

# Performing the hypothesis test

# Calculating the test statistic
test_statistic <- (predicted_price - expected_price) /
  (sqrt(sum((data$sqft - mean(data$sqft))^2) / (length(data$sqft) - 2)))

# Getting the p-value
p <- pnorm(test_statistic, lower.tail = FALSE)
# Setting the significance level (alpha)
a <- 0.01

# Checking if p-value is less than alpha
if (p < a) {
  print("Reject the null hypothesis")
} else {
  print("Fail to reject the null hypothesis")
}
```

```
## [1] "Fail to reject the null hypothesis"
```

```r
#C

#Setting the confidence level
confidence_level <- 0.99
# Defining the house size
house_size <- 1500

# Predicting the price for a house with 1500 square feet
predicted_price <- predict(model_1, newdata = data.frame(sqft = house_size),
                           interval = "confidence", level = confidence_level)

# Extracting lower bound, and upper bound from the prediction
lower <- predicted_price[2]
upper <- predicted_price[3]

# Printing the results
cat("99% Confidence Interval:", lower, "to", upper, "\n")
```

```
## 99% Confidence Interval: 68092 to 97781.84
```

```r
#D FOR 2000

# Filtering the traditional-style houses
traditional_houses <- data %>% filter(Traditional == 1)
# Quadratic regression model
```

```r
model_2 <- lm(price ~ I(sqft^2), data = traditional_houses)

# Defining the house size
house_size <- 2000

# Calculating the marginal effect
marginal_effect <- 2 * summary(model_2)$coefficients["I(sqft^2)", "Estimate"] * house_size

# Null hypothesis value
null_value <- 70

# Performing the hypothesis test

# Calculating the test statistic
test_statistic <- (marginal_effect - null_value) /
  (2 * summary(model_2)$coefficients["I(sqft^2)", "Std. Error"] * house_size)

# Getting the p-value
p <- pnorm(test_statistic, lower.tail = TRUE)
# Set the significance level
a <- 0.01

# Checking if p-value is less than alpha
if (p < a) {
  print("Reject the null hypothesis")
} else {
  print("Fail to reject the null hypothesis")
}
```

```
## [1] "Reject the null hypothesis"
```

```r
#D FOR 3000

# Defining the house size
house_size <- 3000
# Calculating the marginal effect
marginal_effect <- 2 * summary(model_2)$coefficients["I(sqft^2)", "Estimate"] * house_size
# Null hypothesis value
null_value <- 70

# Performing the hypothesis test

# Calculating the test statistic
test_statistic <- (marginal_effect - null_value) /
  (2 * summary(model_2)$coefficients["I(sqft^2)", "Std. Error"] * house_size)

# Getting the p-value
p <- pnorm(test_statistic, lower.tail = TRUE)
# Set the significance level
a <- 0.01

# Checking if p-value is less than alpha
if (p < a) {
```

```r
  print("Reject the null hypothesis")
} else {
  print("Fail to reject the null hypothesis")
}
```

```
## [1] "Fail to reject the null hypothesis"
```

```r
#E FOR 2000

# Fitting a log-linear regression model
model_3 <- lm(log(price) ~ sqft, data = traditional_houses)

# Defining the house size
house_size <- 2000
# Predicting the price for a house with 2000 square feet
predicted_price <- predict(model_3, newdata = data.frame(sqft = house_size))
# Calculate the marginal effect at 2000 square feet
marginal_effect <- predicted_price * coef(model_3)["sqft"]

# Null hypothesis value
null_value <- 70

# Performing a one-tailed hypothesis test

# Calculating the test statistic
test_statistic <- (marginal_effect - null_value) /
  (summary(model_3)$coefficients["sqft", "Std. Error"])

# Getting the p-value
p <- pnorm(test_statistic, lower.tail = TRUE)
# Setting the significance level
a <- 0.01

# Checking if p-value is less than alpha
if (p < a) {
  cat("Reject the null hypothesis")
} else {
  cat("Fail to reject the null hypothesis")
}
```

```
## Reject the null hypothesis
```

```r
#E FOR 3000

# Defining the house size
house_size <- 3000
# Predicting the price for a house with 3000 square feet
predicted_price <- predict(model_3, newdata = data.frame(sqft = house_size))
# Calculate the marginal effect at 3000 square feet
marginal_effect <- predicted_price * coef(model_3)["sqft"]
# Null hypothesis value
null_value <- 70
```

```r
# Performing a one-tailed hypothesis test

# Calculating the test statistic
test_statistic <- (marginal_effect - null_value) /
  (summary(model_3)$coefficients["sqft", "Std. Error"])

# Getting the p-value
p <- pnorm(test_statistic, lower.tail = TRUE)
# Setting the significance level
a <- 0.01

# Checking if p-value is less than alpha
if (p < a) {
  cat("Reject the null hypothesis")
} else {
  cat("Fail to reject the null hypothesis")
}
```

```
## Reject the null hypothesis
```