Math 300 Lesson 34 Notes

Interpreting Regression Tables

YOUR NAME HERE

July, 2022

Contents

Objectives

1. Build and interpret a regression table and use correct terminology.
2. Write the hypotheses from the regression table.
3. Use the p-value from a regression table to make a decision about the relationship between the explanatory and response variables.

Reading

Chapter 10 - 10.2

Lesson

There are no learning checks for this section.

- This is a chance to review regression and also to conduct a hypothesis test.
- Focus on the interpretation of the results.

Libraries

```
library(tidyverse)
library(infer)
library(moderndive)
```

Problem

We will now use the inference tools we have developed in a regression setting. In this lesson we will review and spend some time on interpreting regression output.

Data

Let's get the data again.

```
evals_ch5 <- evals %>%
  select(ID, score, bty_avg, age)
glimpse(evals_ch5)
```

```
## Rows: 463
## Columns: 4
## $ ID      <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,~
## $ score   <dbl> 4.7, 4.1, 3.9, 4.8, 4.6, 4.3, 2.8, 4.1, 3.4, 4.5, 3.8, 4.5, 4.~
## $ bty_avg <dbl> 5.000, 5.000, 5.000, 5.000, 3.000, 3.000, 3.000, 3.333, 3.333,~
## $ age     <int> 36, 36, 36, 36, 59, 59, 59, 51, 51, 40, 40, 40, 40, 40, 40, 40~
```

EDA

The scatterplot revealed what appeared to be a slight positive association between the two variables. But is this due to the randomness associated with sampling?

```
# Complete the code
# ggplot(evals_ch5,
#        aes(x = _________, y = _________)) +
#   geom_point() +
#   labs(x = "Beauty Score",
#        y = "Teaching Score",
#        title = "Relationship between teaching and beauty scores") +
#   geom_smooth(method = "lm", se = FALSE) +
#   theme_classic()
```

Regression Model

Let's build the linear regression model.

```
# Complete the code
# Fit regression model:
# score_model <- lm(________ ~ _________________, data = evals_ch5)
```

```
# Complete the code
# Get regression table:
# get_regression_table(_____________)
```

- Some review:
- How do we interpret the 0.067 value in the `estimate` column?

  For every increase of one unit in "beauty" rating, there is an associated increase, on average, of 0.067 units of evaluation score.

Note that this table has a column labeled `p-value`. This implies that we conducted a hypothesis test. The hypothesis test for the slope is:

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0$$

A p-value is the probability of obtaining a test statistic just as extreme or more extreme than the observed test statistic assuming the null hypothesis H_0 is true.

Conclusion

Answer:

Your turn

In the `moderndive` package is a dataset on electric vehicle charging sessions for a workplace charging program.

We are interested in whether there is an association between the total energy used at a station, `kwh_total`, and the amount paid for a charging session, `dollars`. We are only interested in those workers that paid more than $0.50 for a session. Repeat the analysis from the book on this dataset.

Data

```
ev_chp10 <- ev_charging %>%
  select(kwh_total,dollars) %>%
  filter(dollars>0.5)
```

```
glimpse(ev_chp10)
```

```
## Rows: 184
## Columns: 2
## $ kwh_total <dbl> 6.76, 7.84, 2.76, 4.10, 7.21, 3.11, 9.31, 7.89, 6.30, 5.48, ~
## $ dollars   <dbl> 0.58, 0.92, 1.25, 0.83, 1.58, 1.33, 1.25, 1.33, 1.67, 0.92, ~
```

EDA

Build a scatterplot with energy used on the vertical axis and money paid on the horizontal axis.

```
#ggplot(ev_chp10,
#       aes(y = ________, x = ________)) +
#  geom_point() +
#  labs(x = "Money Paid",
#       y = "Total Energy",
#       title = "Relationship between money paid for charging session and total energy used") +
#  geom_smooth(method = "lm", se = FALSE) +
#  theme_classic()
```

Regression Model

Let's build the linear regression model.

```
# Complete the code
# Fit regression model:
```

```
# Complete the code
# Get regression table:
```

Conclusion

Answer:

Documenting software

- File creation date: 2022-07-15
- R version 4.1.3 (2022-03-10)
- `tidyverse` package version: 1.3.1
- `moderndive` package version: 0.5.4
- `infer` package version: 1.0.2