

AE 11: Inference for Multiple Linear Regression

Credit Cards

! Important

- Open [RStudio](#) and create a subfolder in your AE folder called “AE-11”.
- Go to the [Canvas](#) and locate your AE-11 assignment to get started.
- Upload the `ae-11.qmd` file into the folder you just created. The `.qmd` and PDF responses are due in Canvas. You can check the due date on the Canvas assignment.

Packages + data

```
library(tidyverse)
library(ggformula)
library(broom)
library(knitr)
library(ISLR2)
library(GGally)
library(yardstick)
```

The data for this AE is from the `Credit` data set in the **ISLR2** R package. It is a simulated data set of 400 credit card customers. We will focus on the following variables:

Predictors

- **Income:** Annual income (in 1000’s of US dollars)
- **Rating:** Credit Rating

Response

- **Limit:** Credit limit

Analysis goal

The goals of this activity are to:

- Perform inference for multiple linear regression
- Conduct/interpret hypothesis tests
- Construct/interpret confidence intervals

Exercise 0

Fit and display two linear regression models predicting **Limit** from **Income** and **Rating**. In one, include an interaction term between the two variables and in the other, do not.

Exercise 1

Consider the model *without* an interaction term. Perform a hypothesis test on **Rating** (fill in the blanks where appropriate:

1. **Set hypothesis:** $H_0 : \beta_{Rating}[fillin]$ vs. $\beta_{Rating}[fillin]$. Restate these hypothesis in words: [fill in]
2. **Calculate test statistics and p-value:** The test statistic is [fillin]. The p-value is [fillin].
3. **State the conclusion:** [fill in]

Exercise 2

Consider the model *with* an interaction term. Interpret the p-value associated with **Income** and the p-value associated with the interaction term.

Exercise 3

Generate 95% confidence intervals for the model *without* an interaction term. Hint: use the `tidy` function with the argument `conf.int = TRUE`. Interpret the confidence interval for **Income** and explain why the combination of p-value and confidence interval makes sense.

Exercise 4

What does it mean for two things to be **independent** in statistics (feel free to use google)? Do we think our p-values/confidence intervals are **independent** across variables?

To submit the AE

! Important

- Render the document to produce the PDF with all of your work from today's class.
- Upload your QMD and PDF files to the Canvas assignment.