IPS9 in R: Logistic Regression (Chapter 14)

Margaret Chien and Nicholas Horton (nhorton@amherst.edu)

July 18, 2018

Introduction and background

These documents are intended to help describe how to undertake analyses introduced as examples in the Ninth Edition of *Introduction to the Practice of Statistics* (2017) by Moore, McCabe, and Craig.

More information about the book can be found here. The data used in these documents can be found under Data Sets in the Student Site. This file as well as the associated R Markdown reproducible analysis source file used to create it can be found at https://nhorton.people.amherst.edu/ips9/.

This work leverages initiatives undertaken by Project MOSAIC (http://www.mosaic-web.org), an NSF-funded effort to improve the teaching of statistics, calculus, science and computing in the undergraduate curriculum. In particular, we utilize the `mosaic` package, which was written to simplify the use of R for introductory statistics courses. A short summary of the R needed to teach introductory statistics can be found in the mosaic package vignettes (http://cran.r-project.org/web/packages/mosaic). A paper describing the mosaic approach was published in the *R Journal*: https://journal.r-project.org/archive/2017/RJ-2017-024.

Chapter 14: Logistic Regression

This file replicates the analyses from Chapter 14: Logistic regression.

First, load the packages that will be needed for this document:

```
library(mosaic)
library(readr)
```

Section 14.1: The Logistic Regression Model

Example 14.3: Comparing the proportions of female and make Instagram users

```
Instagram <- read_csv("https://nhorton.people.amherst.edu/ips9/data/chapter14/EG14-03INSTAGR.csv")
```

```
## Parsed with column specification:
## cols(
##   Sex = col_character(),
##   SexNum = col_integer(),
##   User = col_character(),
##   Count = col_integer()
## )
```

```
# For odds ratio, should I use oddsRatio()?
```

Section 14.2: Inference for Logistic Regression