

M/STAT 501: Weighted Least Squares Activity

Professor Ratings

Bleske-Rechek and Fritsch (2011) analyzed a data set of the ratings of 366 instructors at one large campus in the Midwest. Each instructor in the data had at least 10 ratings over a several year period. Students provided ratings from 1 (worst) to 5 (best). These data are built into R in the `alr4` library.

```
library(pracma) # eye function
library(alr4) # data set
data(Rateprof)
```

Let Y_{ij} be the quality rating of the i th instructor by the j th student, $j = 1, \dots, n_i$, and $\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i$ be the mean quality rating for the i th instructor. Similarly, let x_{1ij} and x_{2ij} be the easiness and helpfulness ratings, respectively, of the i th instructor by the j th student, with mean easiness and mean helpfulness for the i th instructor denoted by \bar{x}_{1i} and \bar{x}_{2i} . Note that the data set only reports mean ratings, not individual student's ratings.

Do in class:

1. Assume $E(Y_{ij}|X) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij}$ and $Var(Y_{ij}|X) = \sigma^2$. Derive the expression for $E(\bar{Y}_i|X)$ and $Var(\bar{Y}_i|X)$.
2. If we fit a linear model to $E(\bar{Y}_i|X)$, would it meet the constant variance assumption? Explain why or why not.
3. Let $\mathbf{Y} = (\bar{Y}_1 \ \bar{Y}_2 \ \dots \ \bar{Y}_{366})'$ be the response vector for this data set with variance-covariance matrix $Var(\mathbf{Y}) = \sigma^2 \mathbf{\Omega}$. Write out the elements in the first four rows and first four columns of $\mathbf{\Omega}$, i.e., report the 4×4 matrix that consists of elements in rows 1-4 and columns 1-4.
4. Generate plots to investigate the distributions and relationships between the three variables of interest.

```
# Add code here
```

5. Fit the weighted least squares model. Write the equation of the fitted model, and interpret each of the three fitted coefficients in context of the problem.

```
# Add code here
```

6. Fit the ordinary least squares fit to these data. How do the coefficients change? How does the residual standard error change? Why?

```
# Add code here
```