Notes from Gelman & Hill

Chapter 1

What is multilevel regression?

The example of regressing student test scores by school. Three options:

- A model where coefficients vary by school ($y = \alpha_j + \beta_j x + error$)
- A model with more than variance component (??)
- A regression with many predictors including an indicator variable for each school in the data

Chapter 4

What is a z-score? A z-score is an estimation of the coefficients based on a standardized input calculated as the input minus its mean divided by 2 standard deviations.

Two standard deviations are used instead of one for interpretability of binary outputs. If using one standard deviation to calculate the z-score of a coefficient for a binary input, the interpretation would be on the effect of half difference between the two inputs of X. Whe

Chapter 18

`lm` function in R computes $\beta^{MLE} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}y$ which results in a best fit (least square of errors) for the parameters.

Important conclusion: maximizing likelihood of the outputs given the inputs and the parameters is equivalent to minimizing the sum of squares error term

Weighted least squares (β^{WLS}): Adds a W matrix for weighting each observation as having relatively more or less influence on the optimization:

$$\beta^{WLS} = (\mathbf{X}^T W \mathbf{X})^{-1}\mathbf{X}^T W y$$

Bayesian inference in a sentence: Multiplying a prior distribution reflecting our beliefs by the maximimum likelihood learned from the data forms the posterior distribution, from which random draws reflect both inferences from the data and our beliefs.