STAT 151A Lecture 17

Henry Liev

6 October 2025

Remark 0.1 (Binary Variables)

$x \in \{0, 1\}$

We can relationship between Y and X by just calculating mean of Y when $x = 0$ and mean of Y when $x = 0$

$n_m = \bar{Y}_{male}, n_f = \bar{Y}_{female}$

$X = \begin{pmatrix} \vec{M} & \vec{F} \end{pmatrix}$, $\vec{M}$ binary vector of male or not, $\vec{F}$ binary vector of female or not, project Y onto $\text{Col}(X)$ build $\mathbf{X}\hat{\beta}, \hat{\beta} = (\mathbf{X}^\intercal\mathbf{X})^{-1}\mathbf{X}^\intercal Y$

$\mathbf{X}^\intercal\mathbf{X} = \begin{pmatrix} n_m & 0 \\ 0 & n_f \end{pmatrix}$

$(\mathbf{X}^\intercal\mathbf{X})^{-1} = \begin{pmatrix} \frac{1}{n_m} & 0 \\ 0 & \frac{1}{n_f} \end{pmatrix}$

$(\mathbf{X}^\intercal\mathbf{X})^{-1} = \begin{pmatrix} \bar{Y}_m \\ \bar{Y}_f \end{pmatrix}$

Remark 0.2 (Alternative Design Matrix)

$\begin{pmatrix} \vec{1} & \vec{F} \end{pmatrix}$ New $\hat{\beta}$

$\mathbf{X}^\intercal\mathbf{X} = \begin{pmatrix} n & n_f \\ n_f & n_f \end{pmatrix}$

$(\mathbf{X}^\intercal\mathbf{X})^{-1} = \frac{1}{nn_f - n_f^2} \begin{pmatrix} n_f & -n_f \\ -n_f & n \end{pmatrix}$

$\mathbf{X}^\intercal Y = \begin{pmatrix} \sum y_i \\ \sum_{i \in f} y_i \end{pmatrix}$

$(\mathbf{X}^\intercal\mathbf{X})^{-1}\mathbf{X}^\intercal Y = \begin{pmatrix} \bar{y}_m \\ \bar{y}_f - \bar{y}_m \end{pmatrix}$

Remark 0.3

Baseline: $y \sim 1$, y on intercept $\mathbf{X} = \begin{bmatrix} \vec{1} \end{bmatrix}$

Gender: $y \sim$ gender $\mathbf{X} = \begin{bmatrix} \vec{1} & \vec{F} \end{bmatrix}$

Rate $y \sim$ rate $\mathbf{X} = \begin{bmatrix} \vec{1} & \vec{r} \end{bmatrix}$

Parallel lines model (Gender + rate) $y \sim$ gender + rate $\mathbf{X} = \begin{bmatrix} \vec{1} & \vec{F} & \vec{r} \end{bmatrix}$

Non-parallel lines model (Gender $\times$ rate) $y \sim$ gender $*$ rate $\mathbf{X} = \begin{bmatrix} \vec{1} & \vec{F} & \vec{r} & \vec{r} + \vec{F} \end{bmatrix}$

Principle of marginality: If you included an interaction, include the main effects too