CSSS/POLS 510 MLE Lab

Lab 3. Heteroskedastic Normal

Ramses Llobet

Preview

- Last lab review.
- OLS and MLE connection.
- Maximum Likelihood Estimation: inference.
- Heteroskedastic normal.

1. Last lab review: Least Squares

- Linear homoskedastic:

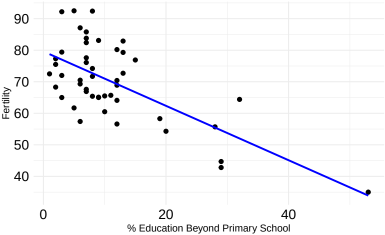
$$\mathbf{y} = \mathbf{X}\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2)$$

- Coefficient estimates:

$$\hat{\beta}_j = \frac{\text{Cov}(X_j, Y) - \sum_{k \neq j} \hat{\beta}_k \text{Cov}(X_j, X_k)}{\text{Var}(X_j)}$$

- Matrix Algebra solution:

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



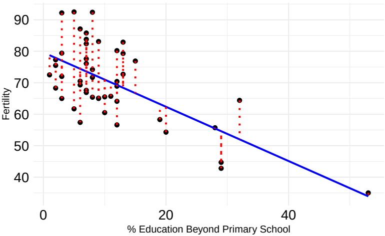
1. Last lab review: Least Squares

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta} \quad (1)$$

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}} \quad (2)$$

$$\sum \hat{e}_i{}^2 = \sum (y_i - \hat{y}_i)^2 \quad (3)$$

$$\sum \hat{e}_i{}^2 = \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \quad (4)$$



Note: (3) and (4) are equivalent, and provide the *Sum of Squared Residuals* (SSR), which is a measurement of fitness.

Choosing the best combination of $\hat{\beta}_j$ that minimizes the SSR provide the best fit for the model.

2. OLS and MLE connection

1. Likelihood function of homoskedastic normal.

$$L(\beta, \sigma^2 \mid \mathbf{y}, \mathbf{X}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_i - \mathbf{X}_i\beta)^2}{2\sigma^2}\right)$$

2. Log-likelihood of homoskedastic normal.

$$\log L(\beta, \sigma^2 \mid \mathbf{y}, \mathbf{X}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \mathbf{X}_i\beta)^2$$

3. Sufficient statistics for estimation.

$$\arg\min_{\beta} \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \mathbf{X}_i\beta)^2$$

2. OLS as Normal homoskedastic model

Two **different** notations for the **same** model.

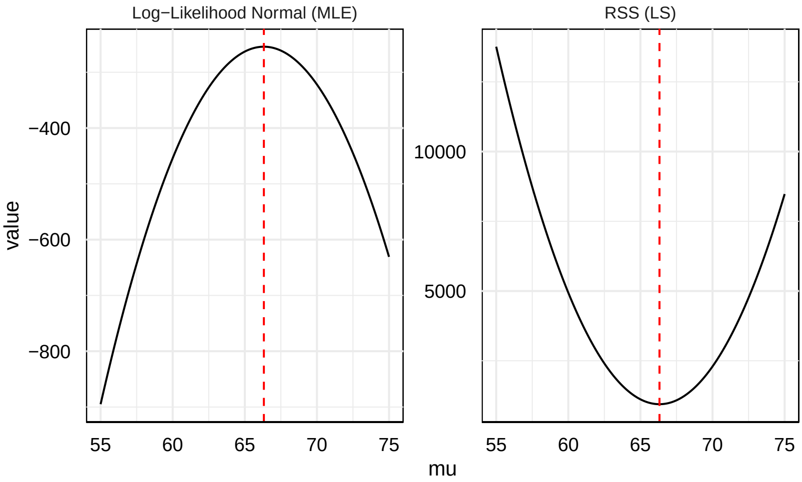
LS notation:

$$\varepsilon \sim N(0, \sigma^2) \quad \text{(\textit{stochastic})}$$
$$Y_i = x_i\beta \quad \text{(\textit{systematic})}$$
$$Y_i = x_i\beta + \varepsilon \quad \text{(\textit{stochastic} + \textit{systematic})}$$

MLE notation:

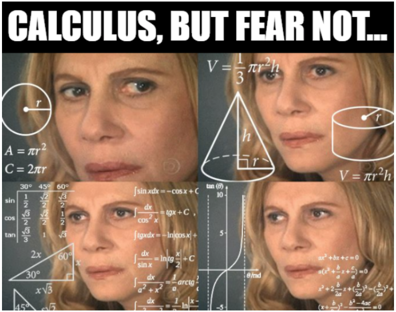
$$Y_i \sim N(\mu_i, \sigma^2) \quad \text{(\textit{stochastic})}$$
$$\mu_i = x_i\beta \quad \text{(\textit{systematic})}$$
$$Y_i \sim N(x_i\beta, \sigma^2) \quad \text{(\textit{stochastic} + \textit{systematic})}$$

2. OLS as Normal homoskedastic model



2. Optimization: analytical approach

- In LS, optimization **minimizes** the **sum of squared residuals** (SSR), while in MLE, it **maximizes** the **log-likelihood** function with respect to the model parameters.
- **Analytical solution**: taking the derivative of the **objective function** with respect to the **parameters**, setting it to zero, and solving for the parameters, if a closed solution exist.



2. Optimization: numerical approach

- Instead of using calculus, we will approach the estimation via numerical methods.
  - Numerical **optimization algorithms** (like gradient descent or Newton-Raphson) are used to **iteratively** update parameter estimates until the **objective function** is optimized.
- `optim()` performs **general-purpose optimization**, helping to minimize or maximize a user-defined function.
  - *Arguments*:
    - **`par`**: Initial parameter values to start the optimization.
    - **`fn`**: The function to be minimized or maximized.
    - **`method`**: Specifies the optimization algorithm (e.g., `"BFGS"`, `"Nelder-Mead"`, `"CG"`).
    - **Hessian**: if `TRUE` returns information matrix for variance-covariance.

3. Maximum Likelihood Estimation

- How do we estimate the MLE?
  1. Define a probability model (PDF): $Y_i \sim N(\mu_i, \sigma^2)$.
  2. Derive the log-likelihood function.
  3. Reduce to sufficient statistics and substitute systematic component.
  4. Use `optim()` or any other function to find the maxima.

4. Heteroskedastic normal

- Steps
  - The full `R` code can be found here from Chris' website
    2.1 Generate Data
    2.2 Fit OLS - `lm()`
    2.3 Fit MLE - `optim()`
- Open the file `Lab3.Rmd` to find the code and contents of today's lab.

F I N