Week 4

Properties of OLS for any sample given assumptions:

- Expected values/unbiasedness under MLR1-MLR4
- Variance formulas under MLR1-MLR5
- Gauss-Markov Theorem under MLR1-MLR5
- Exact sampling distributions under MLR1-MLR6

Why is this important

- To understand what we have to assume to obtain *causal estimates* of our parameters of interest
- And we need variance formulas and sampling distributional assumption to conduct **inference**

Assumption MLR.3 (No Perfect Collinearity)

- No ***perfect collinearity*** in the sample and therefore the population
- None of the independent variables is **constant** and there are **no exact relationships** among the independent variables - you cannot keep regressors fixed to examine the relationship of one *x* variable

Assumption MLR.4 (Zero conditional mean)

- **$E(u_i|x_i) = 0$**
- The value of the explanatory variables must contain no information about the mean of the unobserved factors - more likely to hold in the MLR because fewer things end up in the error

Assumption MLR.5 (Homoskedasticity)

- **$Var(u_i|x_i) = \sigma^2$**
- Error distributions have equal variance (variability) given different x_i

If MLR.1 - MLR.5 Holds:

- **Unbiased estimator of the error variance**
  - $E(\hat{\sigma}^2) = \sigma^2$
- Estimating the error variance:
  - $$\hat{\sigma}^2 = \left(\sum_{i=1}^{n} \hat{u}_i^2\right) \Big/ (n-k-1) = SSR/(n-k-1)$$
  - **DOF** = (*n* - *k* - *1*)

Estimating sampling variance of OLS

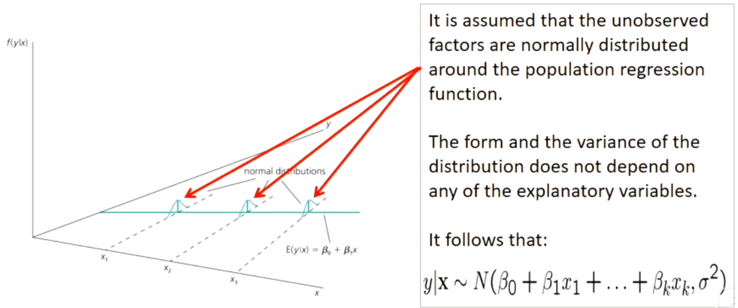
- $$se(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)} = \sqrt{\hat{\sigma}^2 / \left[SST_j(1-R_j^2)\right]}$$
- We get standard error because we do not observe population variance

Gauss-Markov Theorem

- Under assumptions **MLR.1 - MLR.5** - the OLS estimators are the **best linear unbiased estimators (BLUE)** of the regression coefficients

Assumption MLR.6 - Normality of error terms

- Population error (u_i) is independent of the regressor (x_i) and is *__normally distributed__* with zero mean and variance σ^2



Random variable

- Variable that takes on a set of all possible numerical values that are determined probabilistically
    - What is observed is a **realisation** of a random variable
    - Uppercase = random variable
    - Lowercase = realisation of random variable

Discrete Random Variables

- One that takes on only a **finite** number of values

Probability density function (PDF)

- Summarises the possible outcomes of X and the corresponding probabilities

Cumulative distribution function (CDF)

- Sum of PDFs over all values of x_j
- $F(X) = P(X \leq x) = \sum_{x_j < x}$
- Probability that realisation is less than or equal to x

Continuous Random Variables

- There are an infinite number of values that can be taken with each taking zero probability - it only works with ranges
    - $$F(x) = P(X \leq x) = \int_{-\infty}^{\infty} f(x)dx$$