Lecture 7

Normal Errors

Ananya Iyengar

February 2026

St. Stephen's College

This corresponds to Section 3.9 of Gujarati and Porter (2009) and derives from selected portions of Hansen (2014) and Tsun (2020).

1 Finite Sample Analysis

With a finite n, one needs to assume normality of regressions residuals. The mean of this normal distribution is assumed to be zero, and the variance a constant. While testing hypotheses, the assumption of a zero mean helps prove the unbiasedness of the OLS estimator and the assumption of homoskecasticity helps assume a constant estimator variance. In a finite sample, this variance is unknown, which implies using the t-test for hypothesis testing.

Assuming normality ex-ante is not enough – after estimation, one must make sure that regression residuals are indeed normally distributed. Failure to do so negates the use of the t-distribution.

With a finite sample, normality can be checked by:

1. Plotting a histogram of residuals.
2. Plotting a normal probability plot.

2 Moments of Random Variable

Let X be a random variable. Then the k^{th} moment of X is given by $E[X^k]$. If $c \in \mathbb{R}$, then the k^{th} moment of X around c is given by $E[(X-c)^k]$. Then,

1. The first moment of X is the mean $\mu = E[X]$.
2. The second moment of X centered around μ is the variance $\sigma^2 = E[(X-\mu)^2]$.
3. The third moment of X, centered around μ and standardised by σ is given by the skewness $S = E[(\frac{X-\mu}{\sigma})^3]$. Positively skewed distributions are ones with the mean larger than the median; while negatively skewed distributions are ones with the median larger than the mean.
4. The fourth moment of X, centered around μ and standardised by σ, is given by the kurtosis $K = E[(\frac{X-\mu}{\sigma})^4]$. A positive kurtosis is associated with a more peaked distribution.

2.1 Moment Generating Functions

Let X, Y be independent random variables and $a, b \in \mathbb{R}$ be scalars. Then,

1. The moment generating function $M_X : \mathbb{R} \to [0, \infty)$ is defined by $M_X(t) = E[e^{tX}]$.
2. The n^{th} moment of M_X is the n^{th} derivative $M_X^{(n)}$. This is used to derive the moments listed above.
3. $M_{aX+b}(t) = e^{tb} M_X(at)$
4. M_X uniquely identifies a distribution.

3 Asymptotic Results

3.1 The Jarque-Bera Test

The Jarque-Bera test is an asymptotic/large-sample test of normality for data with unknown mean and variance. For a normal distribution, $S = 0$ and $K = 3$. Therefore,

H_0 : Data is from a normal distribution

H_a : Data is not from a normal distribution The test statistic:

$$JB = \frac{n}{6}[S^2 + \frac{(K-3)^2}{4}] \sim_{asy} \chi^2_{(2)}$$

Decision Rule: Reject the null if $\chi^2_{obs,2} > \chi^2_{c,2}$.

3.2 The Central Limit Theorem

Let $X_1, ...X_n$ be a sequence of i.i.d random variables with mean $\mu \equiv E[X]$ and $E[X^2] < \infty$. Then,

$$\text{As } n \to \infty, \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \to N(0,1)$$

This is the most restricted version of the CLT. There are extensions where under certain conditions, normality holds even when the Xs are not identically distributed.

Proof. Without loss of generality, assume $\mu = 0$. This implies $V(X) = E[X^2] = \sigma^2$. Also, let $\bar{Z}_n = \frac{\bar{X}}{\sigma/\sqrt{n}} = \frac{1}{\sigma\sqrt{n}} \sum X_i$.

Let Y be an arbitrary random variable. Then the MGF of Y is given by $M_Y(s)$. Using the 2^{nd} order Taylor expansion around 0,

$M_Y(s) \approx M_Y(0)\frac{s^0}{0!} + M'_Y(0)\frac{s^1}{1!} + M''_Y(0)\frac{s^2}{2!}$
$= 1 + E[Y]s + E[Y^2]\frac{s^2}{2}$ (1)

Now, all Xs are independently and identically distributed. Therefore, $M_{\bar{Z}}(t) = M_{\frac{1}{\sigma\sqrt{n}}\sum X_i}(t) = M_{\sum X}(\frac{t}{\sigma\sqrt{n}}) = [M_X(\frac{t}{\sigma\sqrt{n}})]^n$ (2).

Using (1), putting $Y = X$ and $s = \frac{t}{\sigma\sqrt{n}}$, $M_X(\frac{t}{\sigma\sqrt{n}}) \approx 1 + E[X]\frac{t}{\sigma\sqrt{n}} + E[X^2]\frac{t^2}{2\sigma^2 n} = 1 + \frac{t^2/2}{n}$.

Using (2), $M_{\bar{Z}}(t) = [M_X(\frac{t}{\sigma\sqrt{n}})]^n \approx (1 + \frac{t^2/2}{n})$. As $n \to \infty$, $(1 + \frac{x}{n})^n \to e^x$, and therefore $(1 + \frac{t^2/2}{n}) \to e^{\frac{t^2}{2}}$, which is the MGF of the standard normal distribution. Since the MGF uniquely identifies a distribution, we have proved that $\bar{Z}_n$ is asymptotically normal.

□

The above proof is for a specific linear combination of i.i.d. X_is, but the theorem holds for all linear combinations. Specifically, note that the estimated coefficient $\hat{\beta}$ is *also* a linear combination of Xs. Therefore, we can derive an asymptotic distribution of $\hat{\beta}$ that allows us to undertake hypothesis testing!

Specifically, by the CLT, $\sqrt{n}(\hat{\beta} - \beta) \to N(0, \Sigma)$, where Σ is the variance-covariance matrix of the Xs.