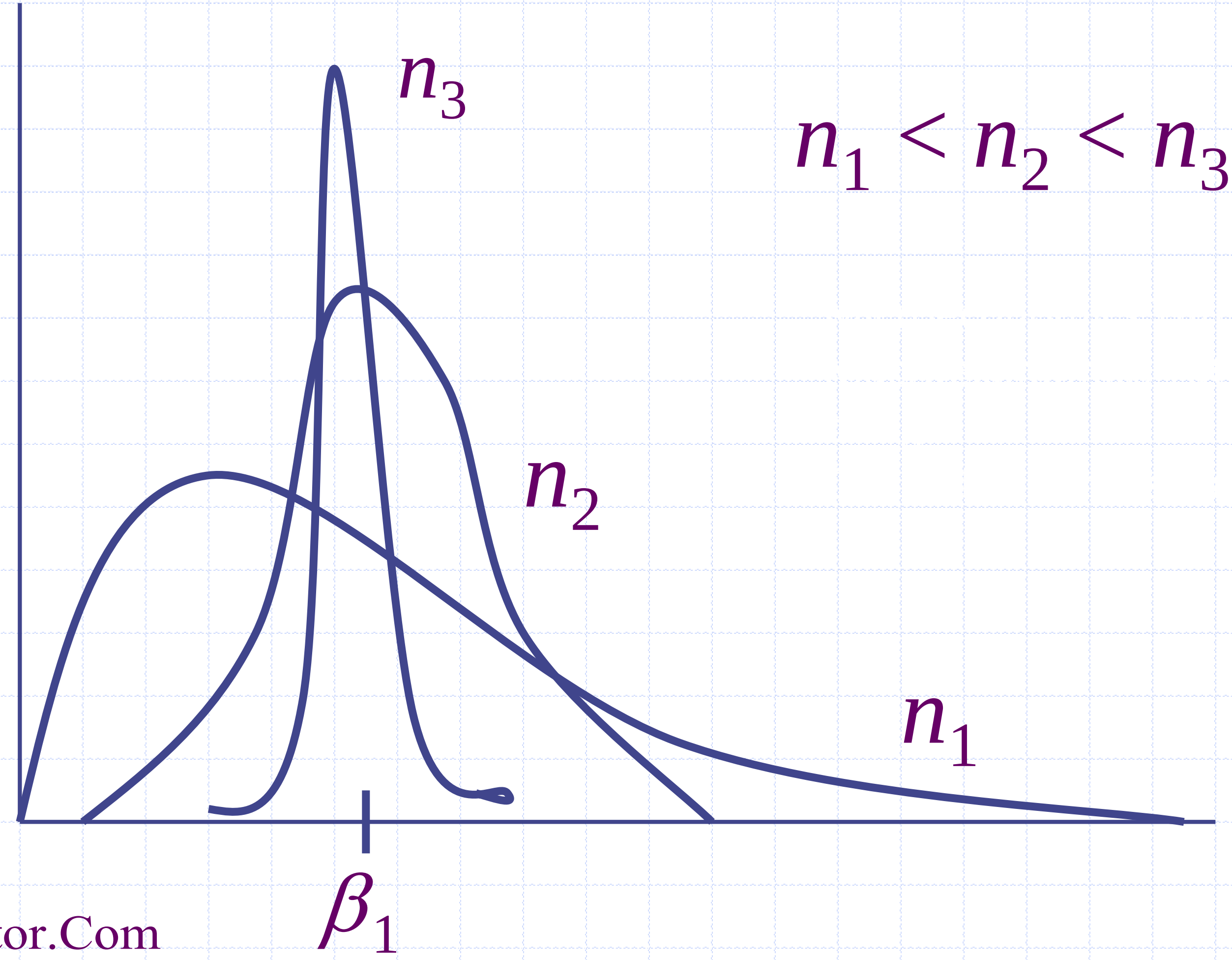
Multiple Regression Analysis

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_k x_k + u$
- 3. Asymptotic Properties

Consistency

- Under the Gauss-Markov assumptions OLS is BLUE, but in other cases it won't always be possible to find unbiased estimators
- In those cases, we may settle for estimators that are consistent, meaning as $n \rightarrow \infty$, the distribution of the estimator collapses to the parameter value

Sampling Distributions as $n \uparrow$



EconTutor.Com

Consistency of OLS

- Under the Gauss-Markov assumptions, the OLS estimator is consistent (and unbiased)
- Consistency can be proved for the simple regression case in a manner similar to the proof of unbiasedness
- Will need to take probability limit (plim) to establish consistency

Proving Consistency

$$\hat{\beta}_1 = \left(\sum (x_{i1} - \bar{x}_1) y_i \right) \Big/ \left(\sum (x_{i1} - \bar{x}_1)^2 \right)$$

$$= \beta_1 + \left(n^{-1} \sum (x_{i1} - \bar{x}_1) u_i \right) \Big/ \left(n^{-1} \sum (x_{i1} - \bar{x}_1)^2 \right)$$

$$\text{plim}\,\hat{\beta}_1 = \beta_1 + Cov(x_1, u) \big/ Var(x_1) = \beta_1$$

$$\text{because } Cov(x_1, u) = 0$$

Large Sample Inference

- Recall that under the CLM assumptions, the sampling distributions are normal, so we could derive t and F distributions for testing
- This exact normality was due to assuming the population error distribution was normal
- This assumption of normal errors implied that the distribution of y, given the x's, was normal as well

Large Sample Inference (cont)

- Easy to come up with examples for which this exact normality assumption will fail
- Any clearly skewed variable, like wages, arrests, savings, etc. can't be normal, since a normal distribution is symmetric
- Normality assumption not needed to conclude OLS is BLUE, only for inference

Central Limit Theorem

- Based on the central limit theorem, we can show that OLS estimators are asymptotically normal
- Asymptotic Normality implies that $P(Z<z)\rightarrow\Phi(z)$ as $n \rightarrow\infty$, or $P(Z<z) \approx \Phi(z)$
- The central limit theorem states that the standardized average of any population with mean μ and variance σ^2 is asymptotically $\sim N(0,1)$, or

$$Z = \frac{\bar{Y} - \mu_Y}{\sigma / \sqrt{n}} \overset{a}{\sim} N(0,1)$$

Asymptotic Normality

Under the Gauss - Markov assumptions,

(i) [redacted] $n\left(\hat{\beta}_j - \beta_j\right) \overset{a}{\sim} \text{Normal}\left(0, \sigma^2\right.$ [redacted] $\left. a_j^2\right)$,

where $a_j^2 = \text{plim}\left(n^{-1} \sum \hat{r}_{ij}^2\right)$

(ii) $\hat{\sigma}^2$ is a consistent estimator of σ^2

(iii) $\left(\hat{\beta}_j - \beta_j\right)$ [redacted] $se\left(\hat{\beta}_j\right) \overset{a}{\sim} \text{Normal}(0,1)$

Asymptotic Normality (cont)

- Because the *t* distribution approaches the normal distribution for large *df*, we can also say that

$$\left(\hat{\beta}_j - \beta_j\right) / se\left(\hat{\beta}_j\right) \overset{a}{\sim} t_{n-k-1}$$

- Note that while we no longer need to assume normality with a large sample, we do still need homoskedasticity