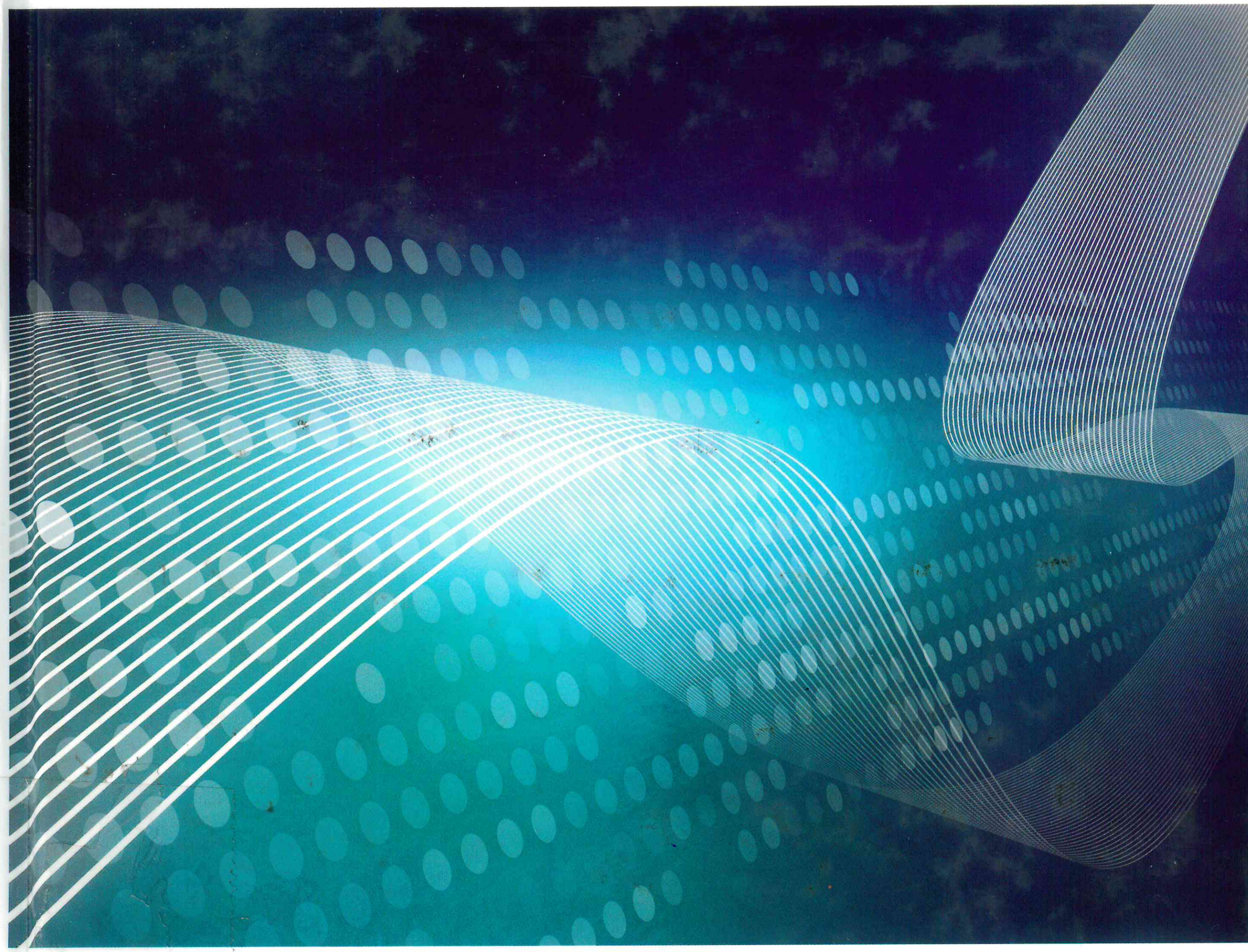
ECONOMETRIC ANALYSIS

SEVENTH EDITION



WILLIAM H. GREENE

in the brackets converges to 0. That leaves

$$\overline{\varepsilon^2} = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i^2.$$

This is a narrow case in which the random variables ε_i^2 are independent with the same finite mean σ^2, so not much is required to get the mean to converge almost surely to $\sigma^2 = E[\varepsilon_i^2]$. By the Markov theorem (D.8), what is needed is for $E[|\,\varepsilon_i^2\,|^{1+\delta}]$ to be finite, so the minimal assumption thus far is that ε_i have finite moments up to slightly greater than 2. Indeed, if we further assume that every ε_i has the same distribution, then by the Khinchine theorem (D.5) or the corollary to D8, finite moments (of ε_i) up to 2 is sufficient. **Mean square convergence** would require $E[\varepsilon_i^4] = \phi_\varepsilon < \infty$. Then the terms in the sum are independent, with mean σ^2 and variance $\phi_\varepsilon - \sigma^4$. So, under fairly weak conditions, the first term in brackets converges in probability to σ^2, which gives our result,

$$\text{plim } s^2 = \sigma^2,$$

and, by the product rule,

$$\text{plim } s^2(\mathbf{X}'\mathbf{X}/n)^{-1} = \sigma^2\mathbf{Q}^{-1}.$$

The appropriate *estimator* of the asymptotic covariance matrix of $\mathbf{b}$ is

$$\text{Est. Asy. Var}[\mathbf{b}] = s^2(\mathbf{X}'\mathbf{X})^{-1}.$$

4.4.4 ASYMPTOTIC DISTRIBUTION OF A FUNCTION OF b: THE DELTA METHOD

We can extend Theorem D.22 to functions of the least squares estimator. Let $\mathbf{f}(\mathbf{b})$ be a set of J continuous, linear, or nonlinear and continuously differentiable functions of the least squares estimator, and let

$$\mathbf{C}(\mathbf{b}) = \frac{\partial \mathbf{f}(\mathbf{b})}{\partial \mathbf{b}'},$$

where $\mathbf{C}$ is the $J \times K$ matrix whose jth row is the vector of derivatives of the jth function with respect to $\mathbf{b}'$. By the Slutsky theorem (D.12),

$$\text{plim } \mathbf{f}(\mathbf{b}) = \mathbf{f}(\boldsymbol{\beta})$$

and

$$\text{plim } \mathbf{C}(\mathbf{b}) = \frac{\partial \mathbf{f}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} = \boldsymbol{\Gamma}.$$

Using a linear Taylor series approach, we expand this set of functions in the approximation

$$\mathbf{f}(\mathbf{b}) = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\Gamma} \times (\mathbf{b} - \boldsymbol{\beta}) + \text{higher-order terms}.$$

The higher-order terms become negligible in large samples if plim $\mathbf{b} = \boldsymbol{\beta}$. Then, the asymptotic distribution of the function on the left-hand side is the same as that on the right. Thus, the mean of the asymptotic distribution is plim $\mathbf{f}(\mathbf{b}) = \mathbf{f}(\boldsymbol{\beta})$, and the asymptotic covariance matrix is $\{\boldsymbol{\Gamma}[\text{Asy. Var}(\mathbf{b} - \boldsymbol{\beta})]\boldsymbol{\Gamma}'\}$, which gives us the following theorem:

THEOREM 4.5 Asymptotic Distribution of a Function of b
If $\mathbf{f}(\mathbf{b})$ *is a set of continuous and continuously differentiable functions of* $\mathbf{b}$ *such that* $\boldsymbol{\Gamma} = \partial\mathbf{f}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}'$ *and if Theorem 4.4 holds, then*

$$\mathbf{f}(\mathbf{b}) \overset{a}{\sim} N\left[\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Gamma}\left(\frac{\sigma^2}{n}\mathbf{Q}^{-1}\right)\boldsymbol{\Gamma}'\right]. \tag{4-36}$$

In practice, the estimator of the asymptotic covariance matrix would be

$$\text{Est. Asy. Var}[\mathbf{f}(\mathbf{b})] = \mathbf{C}[s^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{C}'.$$

If any of the functions are nonlinear, then the property of unbiasedness that holds for $\mathbf{b}$ may not carry over to $\mathbf{f}(\mathbf{b})$. Nonetheless, it follows from (4-25) that $\mathbf{f}(\mathbf{b})$ is a consistent estimator of $\mathbf{f}(\boldsymbol{\beta})$, and the asymptotic covariance matrix is readily available.

Example 4.4 Nonlinear Functions of Parameters: The Delta Method
A dynamic version of the demand for gasoline model in Example 2.3 would be used to separate the short- and long-term impacts of changes in income and prices. The model would be

$$\begin{aligned}\ln(G/Pop)_t = {} & \beta_1 + \beta_2 \ln P_{G,t} + \beta_3 \ln(Income/Pop)_t + \beta_4 \ln P_{nc,t} \\ & + \beta_5 \ln P_{uc,t} + \gamma \ln(G/Pop)_{t-1} + \varepsilon_t,\end{aligned}$$

where P_{nc} and P_{uc} are price indexes for new and used cars. In this model, the short-run price and income elasticities are β_2 and β_3. The long-run elasticities are $\phi_2 = \beta_2/(1-\gamma)$ and $\phi_3 = \beta_3/(1-\gamma)$, respectively. To estimate the long-run elasticities, we will estimate the parameters by least squares and then compute these two nonlinear functions of the estimates. We can use the delta method to estimate the standard errors.

Least squares estimates of the model parameters with standard errors and t ratios are given in Table 4.3. The estimated short-run elasticities are the estimates given in the table. The two estimated long-run elasticities are $f_2 = b_2/(1-c) = -0.069532/(1-0.830971) = -0.411358$ and $f_3 = 0.164047/(1-0.830971) = 0.970522$. To compute the estimates of the standard errors, we need the partial derivatives of these functions with respect to the six parameters in the model:

$$\mathbf{g}_2' = \partial\phi_2/\partial\boldsymbol{\beta}' = [0, 1/(1-\gamma), 0, 0, 0, \beta_2/(1-\gamma)^2] = [0, 5.91613, 0, 0, 0, -2.43365],$$
$$\mathbf{g}_3' = \partial\phi_3/\partial\boldsymbol{\beta}' = [0, 0, 1/(1-\gamma), 0, 0, \beta_3/(1-\gamma)^2] = [0, 0, 5.91613, 0, 0, 5.74174].$$

Using (4-36), we can now compute the estimates of the asymptotic variances for the two estimated long-run elasticities by computing $\mathbf{g}_2'[s^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{g}_2$ and $\mathbf{g}_3'[s^2(\mathbf{X}'\mathbf{X})^{-1}]\mathbf{g}_3$. The results are 0.023194 and 0.0263692, respectively. The two asymptotic standard errors are the square roots, 0.152296 and 0.162386.

4.4.5 ASYMPTOTIC EFFICIENCY

We have not established any large-sample counterpart to the Gauss–Markov theorem. That is, it remains to establish whether the large-sample properties of the least squares estimator are optimal by any measure. The Gauss–Markov theorem establishes finite