Econ 203B: Single Equation Models

Mean Squared Error

Michael Powell
Department of Economics, UCLA
February 4th, 2006

For a finite sample estimator, we are generally interested in two criteria: unbiasedness and efficiency. There is generally a trade-off betwen achieving unbiasedness and reducing the variance of an estimator, however. As a result, we often seek to minimize the mean squared error of an estimator, which is a function of both the bias of an estimator and its variance.

Let $\hat{\theta}$ be an estimator for θ_0. Then we define the mean squared error of the estimator as:

$$MSE\left(\hat{\theta}\right) = E\left[\left(\hat{\theta} - \theta_0\right)^2\right]$$

Proposition 1 $MSE\left(\hat{\theta}\right) = Var\left(\hat{\theta}\right) + \left(bias\left(\hat{\theta}\right)\right)^2$

Proof.

$$\begin{aligned}
MSE\left(\hat{\theta}\right) &= E\left[\left(\hat{\theta} - \theta_0\right)^2\right] \\
&= E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right] + E\left[\hat{\theta}\right] - \theta_0\right)^2\right] \\
&= E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right)^2\right] + E\left[\left(E\left[\hat{\theta}\right] - \theta_0\right)^2\right] \\
&\quad +2E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right)\left(E\left[\hat{\theta}\right] - \theta_0\right)\right] \\
&= Var\left(\hat{\theta}\right) + E\left(bias\left(\hat{\theta}\right)^2\right) \\
&\quad +2E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right)\left(E\left[\hat{\theta}\right] - \theta_0\right)\right] \\
&= Var\left(\hat{\theta}\right) + \left(bias\left(\hat{\theta}\right)\right)^2 \\
&\quad +2E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right)\left(E\left[\hat{\theta}\right] - \theta_0\right)\right]
\end{aligned}$$

Looking at the last term, and applying the law of iterated expectations, we have:

$$\begin{aligned}
2E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right)\left(E\left[\hat{\theta}\right] - \theta_0\right)\right] &= 2E\left[E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right)\left(E\left[\hat{\theta}\right] - \theta_0\right)\right] \middle| \hat{\theta}\right] \\
&= 2E\left[\left(\hat{\theta} - E\left[\hat{\theta}\right]\right) E\left[E\left[\hat{\theta}\right] - \theta_0 \middle| \hat{\theta}\right]\right] \\
&= 2\left[E\left[\hat{\theta}\right] - E\left[\hat{\theta}\right]\right] E\left[E\left[E\left[\hat{\theta}\right] - \theta_0 \middle| \hat{\theta}\right]\right] \\
&= 0
\end{aligned}$$

Putting this together gives us:

$$MSE\left(\hat{\theta}\right) = Var\left(\hat{\theta}\right) + \left(bias\left(\hat{\theta}\right)\right)^2$$

Which is the desired result. ■