Statistics 215

D.A. Freedman

Spring 2008

Notes on ratio estimators and the delta-method

Let (X_i, Y_i) be IID pairs of positive random variables, each variable having several moments. Let

$$R = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i},$$
(1)

a ratio estimator. We seek the asymptotic mean and variance of *R*. Let $\xi_i = [X_i - E(X_i)]/E(X_i)$ and $\eta_i = [Y_i - E(Y_i)]/E(Y_i)$, so that $E(\xi_i) = E(\eta_i) = 0$, while

$$R = \frac{E(Y_i)}{E(X_i)} \frac{1+\eta}{1+\overline{\xi}}$$
(2)

where the overline denotes sample average. For instance,

$$\overline{\eta} = \frac{1}{n} \sum_{i=1}^{n} \eta_i.$$
(3)

Thus,

$$E(R) = \frac{E(Y_i)}{E(X_i)} E\left(\frac{1+\bar{\eta}}{1+\bar{\xi}}\right)$$
(4)

and

$$\operatorname{var}(R) = \frac{E(Y_i)^2}{E(X_i)^2} \operatorname{var}\left(\frac{1+\overline{\eta}}{1+\overline{\xi}}\right).$$
(5)

Technically, of course, *R* may have an infinite variance, or even an infinite mean, e.g., the denominator might have a positive density near 0. We proceed informally, in this respect among others. For *n* large, $\xi \doteq \eta \doteq 0$. Thus,

$$\frac{1+\eta}{1+\overline{\xi}} \doteq 1+\overline{\eta-\xi}.$$
(6)

This is a one-term Taylor series expansion, called the "delta-method"; for rigor, the remainder term would have to be bounded. Now

$$E(\overline{\eta - \xi}) = 0 \tag{7}$$

while

$$\operatorname{var}\left(1+\overline{\eta-\xi}\right) = \frac{1}{n}\operatorname{var}(\eta_i - \xi_i). \tag{8}$$

Thus, R is asymptotically unbiased for $E(Y_i)/E(X_i)$. The asymptotic variance is

$$\frac{1}{n}\frac{E(Y_i)^2}{E(X_i)^2}\operatorname{var}(\eta_i - \xi_i).$$
(9)

To operationalize (9), we can estimate $E(Y_i)/E(X_i)$ by R, and $var(\eta_i - \xi_i)$ by

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i-\overline{Y}}{\overline{Y}}-\frac{X_i-\overline{X}}{\overline{X}}\right)^2.$$
(10)

We turn now to asymptotic bias. This requires an additional term in the expansion:

$$\frac{1+\eta}{1+\overline{\xi}} \doteq (1+\overline{\eta})(1-\overline{\xi}+\overline{\xi}^2) \doteq 1+\overline{\eta-\xi}+\overline{\xi}^2-\overline{\eta\xi}.$$
(11)

The last two terms in the display are responsible for the asymptotic bias:

$$E\left(\overline{\xi}^{2}\right) = \operatorname{var}\left(\overline{\xi}\right) = \frac{1}{n} \frac{\operatorname{var}(X_{i})}{[E(X_{i})]^{2}}$$
(12)

$$E\left(\overline{\eta\xi}\right) = \frac{1}{n} \frac{\operatorname{cov}(X_i, Y_i)}{E(X_i)E(Y_i)}.$$
(13)

The asymptotic bias is therefore

$$\frac{1}{n} \frac{E(Y_i)}{E(X_i)} \left(\frac{\operatorname{var}(X_i)}{[E(X_i)]^2} - \frac{\operatorname{cov}(X_i, Y_i)}{E(X_i)E(Y_i)} \right).$$
(14)

As before, $E(X_i)$ can be estimated by the sample mean of the X_i 's, while $var(X_i)$ is estimated by the sample variance, $cov(X_i, Y_i)$ by the sample covariance, and so forth.