COWLES COMMISSION DISCUSSION PAPER: STATISTICS NO. 363A

The Estimation of Simultaneous Linear Economic Relations

by Tjalling C. Koopmans and William C. Hood

3. Desirable Properties of Estimates

NOTE: Please substitute this for Section 3 of CCDP 363.

In vague terms, the requirement that we shall impose on our estimates is that their deviations from the true parameter values be in some average sense at least as small as those of alternative estimates, at least in large samples. The lack of preciseness in this statement may be remedied by offering definitions of four large sample properties of estimates which will be regarded as desirable.[23] All of these properties relate to the sampling distribution of the estimates. In the present chapter, therefore, these properties must be interpreted with reference to the notion of (imaginary) repeated samples, described in section 1.3, in which the values of the exogenous variables remain the same,[24] while the values of the disturbances are other random drawings from the joint distribution of the disturbances.

A. Consistency. An estimate of a parameter is said to be consistent if, in the sampling distribution of that estimate, the probability of the

---

23. These properties of statistics are discussed in many text books. See for example, H. Cramer, "Mathematical Methods of Statistics"; A.M. Mood, "Introduction to the Theory of Statistics"; S.S. Wilks, "Mathematical Statistics."

24. In Chapter VII, Section 5, this assumption is weakened to the existence of certain probability limits for certain moments of exogenous variables.

absolute value of the discrepancy between the estimate and the true parameter value being less than any given arbitrarily small positive quantity approaches unity as the size of the sample approaches infinity. In symbols, if $P\{E\}$ denotes the probability of an event E, an estimate h_T of a parameter θ is consistent if for all θ

(3.1) $$\lim_{T \to \infty} P\left\{ |h_T - \theta| < \epsilon \right\} = 1$$

where T is the sample size, and ϵ is any positive number, however small. If h_T has this property, it is said to possess the probability limit θ and this relationship of h_T to θ is also denoted

(3.2) $$\underset{T \to \infty}{\text{plim}}\ h_T = \theta.$$

B. Asymptotic normality. A statistic h_T is said to be asymptotically normally distributed if there exist two sequences of numbers η_T and σ_T (where $\sigma_T > 0$) such that, the following limits exist

(3.3) $$\lim_{T \to \infty} \eta_T = \eta, \quad \lim_{T \to \infty} \sigma_T = \sigma,$$

where $\sigma > 0$, and such that, for every λ_1 and λ_2,

(3.4) $$\lim_{T \to \infty} P\left(\eta_T + \frac{\sigma_T}{\sqrt{T}} \lambda_1 \leq h_T \leq \eta_T + \frac{\sigma_T}{\sqrt{T}} \lambda_2 \right) = \int_{\lambda_1}^{\lambda_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx.$$

This says that the probability distribution of h_T approaches more and more closely a normal distribution with mean η_T and standard deviation $\frac{\sigma_T}{\sqrt{T}}$ as T becomes larger and larger, and that the mean η_T and the quantity σ_T associated with that normal distribution approach finite limits. If h_T is asymptotically normal with $\eta = \theta$, then h_T is a consistent estimate of θ.

Asymptotic normality, while not of great importance from the point of view of the purpose of estimation, is a very convenient property in an estimate. Besides making available an asymptotically correct table of percentile points, this property has mathematical advantages, one of which is that it simplifies the definition of the important property of asymptotic efficiency. All the estimates we shall study have the property of asymptotic normality under the assumptions of the present chapter, and retain that property in most of the cases considered in Chapter VII.

C. Asymptotic efficiency. We shall define this as a property possessed by a consistent and asymptotically normal estimate in comparison with all other consistent and asymptotically normal estimates.

An asymptotically normal estimate h_T (characterized by η and defined above) is said to be an asymptotically efficient estimate of a parameter θ if it is consistent ($\eta = \theta$), and if, for any other asymptotically normal and consistent estimate (characterized by $\eta' = \theta$ and σ') we have

(3.5) $$\sigma' \geq \sigma .$$

This says that asymptotically no rival estimate in the category of comparison has a smaller standard deviation. This, of course, is a very desirable property for an estimate to have, if the sample size is such as to give a reasonable approximation to the asymptotic distribution.

If an estimate h_T lacks the property of consistency, this may be because it does not possess a probability limit, or because it possesses a probability limit η which differs from the parameter θ. Only the second possibility is open if h_T is asymptotically normal. In that case, we call h_T asymptotically biased, and $\eta - \theta$ its asymptotic bias.

In the light of the foregoing statements and distinctions, we may reasonably ask why asymptotic efficiency is not the one large-sample property sought. The answer lies in the nature of the compromises which have to be made in the situations which the statistician faces. Illustrations of such compromises are the following. To have asymptotically efficient estimates may sometimes require a more costly estimation procedure than is possible or desirable in the circumstances, or may require solution of mathematical problems so far too difficult to handle. The attainment of asymptotic efficiency may also depend on information which is not available. In particular, the asymptotic efficiency of certain estimates may depend on assumptions concerning the distribution function of the population which in some situations one is not in a position to make. Thus some methods of estimation will be considered, particularly in Chapter VII, under circumstances where they do not yield efficient estimates even asymptotically.

We now mention two properties of estimates for samples of given finite size of which the second is particularly desirable, but which we know how to attain only under rather restrictive assumptions as to the model. These properties are

1. <u>Unbiasedness</u>. An estimate h_T is unbiased for a sample of size T (as distinct from asymptotically unbiased) if its expectation in such samples equals the true parameter value

$$\mathcal{E} h_T = \theta$$

2. <u>Efficiency</u>. An estimate h_T is efficient in a sample of size T (as distinct from asymptotically efficient) if the ratio of its variance (mean square difference about θ from θ) in such samples to the variance about θ of any other estimate h'_T of the same parameter is not greater than one, i.e., if

$$\frac{\mathcal{E}(h_T - \theta)^2}{\mathcal{E}(h'_T - \theta)^2} \leqq 1.$$

Sometimes a property analogous to efficiency can more easily be attained with reference to a more limited class of estimates. The property of being best linear unbiased, defined in section 4.1 below, is of that type.

It must be stated at the outset that disappointingly little is known about the small sample properties of the estimates considered in this Chapter, apart from those simple cases where unbiasedness, or best linear unbiasedness can be proved. Studies by Hurwicz[26] and by Leipnik[27] of simple one-equation models with a lagged endogenous variable create a presumption of considerable bias, in samples of moderate size, whenever lagged endogenous variables are present.

---

26. L. Hurwicz, "Least Squares Bias in Times Series," Article XV in Cowles Commission Monograph 10.

27. R. B. Leipnik, "Distribution of the Serial Correlation Coefficient in a Circularly Correlated Universe," The Annals of Mathematical Statistics, Vol. XVIII, No. 1, March, 1947. (Reprinted as Cowles Commission Paper, New Series, No. 21).