Sequential Probability Ratio Test for Detecting a Unit Root

By K. NAGAI, Y. NISHIYAMA, AND K. HITOMI
Yokohama National University, Kyoto University
and Kyoto Institute of Technology

Consider a scalar AR(1) process;

$$x_t = \beta x_{t-1} + \epsilon_t, \tag{1}$$

where $\{\epsilon_t\}$ are iid disturbances. When $\beta = \pm 1$, the series is said to possess a unit root. Tests for the existence of unit roots in economic time series have been one of the main issues of interest in econometrics since the middle of 1980's. In practice, econometricians focus on testing the null hypothesis of $\beta = 1$ against the alternative of $|\beta| < 1$.

Suppose $\{x_i\}_{i=0}^{\infty}$ with $x_0 = 0$ is generated from (1). When $\beta = 1$, the process is called a unit root process and its behaviour is very different from ones with $|\beta| < 1$. Some macroeconomic time series are said to have a unit root based on the results from DF test. We first briefly review the DF t-statistic. Given a sample $\{x_0, x_1, \dots, x_T\}$, let the OLS estimator of β and an estimator of σ^2 be

$$\hat{\beta}_T = \left(\sum_{i=1}^{T} x_{i-1}^2\right)^{-1} \sum_{i=1}^{T} x_{i-1} x_i, \quad \hat{\sigma}_T^2 = \frac{1}{T}\sum_{i=1}^{T}(x_i - \hat{\beta} x_{i-1})^2.$$

Also denote $W(s)$ as a standard Brownian motion. Then, as $T \to \infty$, we have the following asymptotic results:

$$\frac{\hat{\beta_T} - \beta}{\hat{\sigma_T} / \sqrt{\sum_{i=1}^{T} x_{i-1}^2}} \xrightarrow{\mathrm{d}} \begin{cases} N(0,1) & \text{if } |\beta| < 1 \\ \frac{\int_0^1 W(s) dW(s)}{\sqrt{\int_0^1 W(s)^2 ds}} & \text{if } |\beta| = 1 \\ \text{Cauchy dist.} & \text{otherwise.} \end{cases}$$

Therefore to test the null of unit root against the alternative of stationarity, we use the t value $(\hat{\beta_T} - 1)/\left\{\hat{\sigma}_T / \sqrt{\sum_{i=1}^{T} x_{i-1}^2}\right\}$ which converges to the functional of $W(s)$ above under the null, while it explodes under the alternative. Table 1 shows the size and power of the test by simulation. The size seems to be acceptable, but the power is unsatisfactory for sample sizes of T=50˜150.

Sequential analysis was originally considered by Wald (1947). The idea is as follows. Suppose we can obtain one observation a day, say. We sample every day and when we accumulate "sufficient" information, then we stop sampling and make a statistical decision (estimation or testing). The time when we stop sampling is called the *stopping time*. "How sufficient" is determined by the researchers through some user-determined critical value which controls the accuracy of the results and the expected stopping time. The accuracy is measured by the standard error for estimation and the power Typically, we are better off if we can obtain conclusions earlier due to some cost of sampling or taking time. There exists a trade-off between accuracy and cost of sampling.

Table 1: Rejection rate of DF t test (nominal size=5%, initial value $x_0 = 0$)

	Beta=1.0	Beta=0.95
Rejection rate (T=50)	0.0453	0.1296
Rejection rate (T=100)	0.0460	0.3103
Rejection rate (T=150)	0.0473	0.5580

Table 2: Properties of SURT (size=5%)

$c = 600$	H_0: $\beta = 1.0$	H_1: $\beta = 0.95$
Rejection rate	0.0503	0.3376
E(Tc)	49.644	81.716
E(beta)	0.999	0.950
std(Tc)	25.302	35.894
std(beta)	0.0415	0.0419

LS investigate the statistical properties of the sequential estimator of the AR(1) parameter in model (1). Formally, for a predetermined critical value c, their stopping rule is defined as

$$T_c = \inf\{t > 0| \sum_{i=1}^{t} \frac{x_{i-1}^2}{\hat{\sigma}^2} \geq c\}$$

where

$$\hat{\beta}_T = \left(\sum_{i=1}^{T} x_{i-1}^2 \right)^{-1} \sum_{i=1}^{T} x_{i-1} x_i, \quad \hat{\sigma}_T^2 = \frac{1}{T} \sum_{i=1}^{T} (x_i - \hat{\beta} x_{i-1})^2.$$

We stop sampling when the estimated information $I_t = \sum_{i=1}^{t} x_{i-1}^2 / \hat{\sigma}_t^2$ exceeds a certain predetermined value c, which controls the accuracy of estimation through the "sample size". We write the stopping time as T_c to emphasize that it depends on the choice of c. c controls for the accuracy of the estimation in the sense that the variance of the estimator, $I_{T_c}^{-1}$, is guaranteed to be smaller than c^{-1}. There exists a trade-off between the accuracy of estimation and the cost of observations. If we set c large, T_c will tend to be also large by construction, which will yield a more accurate estimate. If we set c small, sampling will stop relatively earlier, but the accuracy will be lower. Note that T_c itself is a statistic depending on the observations.

Using this stopping time, we calculate sequential estimators by

$$\hat{\beta}_{T_c} = \left(\sum_{i=1}^{T_c} x_{i-1}^2 \right)^{-1} \sum_{i=1}^{T_c} x_{i-1} x_i, \quad \hat{\sigma}_{T_c}^2 = \frac{1}{T_c} \sum_{i=1}^{T_c} (x_i - \hat{\beta} x_{i-1})^2.$$

LS prove the asymptotic normality of $\hat{\beta}_{T_c}$ in the case of $|\beta| \leq 1$:

$$\sqrt{I_{T_c}} (\hat{\beta}_{T_c} - \beta) \xrightarrow{\mathrm{d}} N(0, 1).$$

Further, Shiryaev and Spokoiny (1997) obtain the same result in the explosive case of $|\beta| > 1$ under the assumption of normal disturbances. We can directly apply this result for unit root test which we call a sequential unit root test (SURT).

Table 2 shows Monte Carlo results of SURT for $c = 600$ c controls for the accuracy of inference through the stopping time. We set $c = 600$ so that the average stopping time under the null is about 50. We will show a theoretical relationship between c and the average stopping time (sample size) later. Comparing the size results in Table 2 with those in Table 1, they are mostly satisfactory. In comparing the power, we need to be careful. The SURT procedure requires more sample sizes to stop sampling under the alternative, thus we cannot directly compare them with figures under the null. One point we can make is that DF test under the standard (fixed) sampling, we cannot conclude a unit root exists even if the null is not rejected from a sample of small or medium size (though it seems to prevail in economic literature). Under sequential sampling, however, researchers will be automatically forced to wait until a "sufficient" amount of information is accumulated both under the null and alternative hypotheses.