# ROBUST UNIT ROOT TESTS FOR SEASONAL AUTOREGRESSIVE PROCESS

Yujin Oh* and Beongsoo So†

**Abstract**

The stationarity is one of the most important properties of a time series. We propose robust sign tests for seasonal autoregressive process to determine whether or not a time series is stationary. The tests have an exact binomial null distribution and are robust to the outliers and the heteroscedastic errors. Monte-Carlo simulation shows that the sign test is locally more powerful than the OLSE-based tests for heavy-tailed and/or heteroscedastic error distributions.

KEY WORDS : Robustness, Seasonal unit root test, Nonstationarity

## 1 Introduction

We consider the problem of testing of the random walk hypothesis for seasonal time series data. There have been several research on this subject such as Dickey, Hasza, and Fuller(1984), Hylleberg, Engle, Granger, and Yoo(1990), So and Shin(1999), and So(2001). The usual OLSE-based tests suffer from the size distortion by outliers and the normal assumption may be easily violated in practice. Thus we propose tests which are robust to outliers and valid under weaker assumption that includes any symmetric disturbances. Campbell and Dufour(1995) first used this method for testing a random walk hypothesis. So and Shin(2001) extended the sign test to a mean model and showed several properties of the statistics such as the exact binomial null distribution, the consistency, the robustness to heavy-tailed errors, and the invariance. In this paper we extend the sign test to a seasonal model. The sign tests follow a binomial null distribution so it leads to an exact test. The asymptotic distribution of the test statistics is a normal regardless of the period of seasonality, thus separate tabulations of critical values are not required.

*Department of Statistics, Ewha Womans University, Seoul, 120-750, Korea
†Department of Statistics, Ewha Womans University, Seoul, 120-750, Korea

## 2 The Test Statistics

Consider the seasonal first-order autoregressive, AR(1), model

$$\begin{cases} y_t = \mu_t + u_t \\ u_t = \rho u_{t-d} + e_t, \quad t = 1, \dots, n \end{cases} \tag{1}$$

where $\{y_t\}$ is observation at time $t$, $\rho$ is the unknown parameter, $y_{-d+1}, y_{-d+2}$, $\dots$, $y_0$ are the initial conditions, $d$ is the period of seasonality, and $\mu_t = \mu_{t+d}, t = 1, \dots, n-d$. In this model if $d$ is 4, the time series represents a quarterly data. And if $d$ is 12, the time series represents a monthly data. We are interested in the following hypothesis :

$$H_0 : \rho = 1 \qquad vs. \qquad H_1 : \rho < 1$$

that the null hypothesis presents a seasonal random walk, that is, the time series has a seasonal unit root. We also suppose the model (1) satisfies the following Assumption 1.

**Assumption 1.** $\{e_t\}$ is a sequence of errors with zero conditional median and has no atom at zero given $F_{t-d}$.

Directly from Assumption 1, we have

$$E[sign(e_t)|F_{t-d}] = 0 \qquad \text{and} \qquad P[e_t = 0|F_{t-d}] = 0 \tag{2}$$

where $sign(e_t) = \begin{cases} 1, & e_t \geq 0 \\ -1, & e_t < 0 \end{cases}$. The test statistic we propose is

$$S_{d,n} = \sum_{t=1}^{n} sign(y_t - y_{t-d}) sign(y_{t-d} - \hat{\mu}_{t-d}) \tag{3}$$

where $\hat{\mu}_t$ is the median of $y_1, \dots, y_t$ for the mean model and $F_t$-measurable with no atoms at zero. We consider two sign tests by the types of the mean adjustment: the common mean type, $S^s_{d,n}$, and the seasonal mean type, $S^c_{d,n}$. Since both common and seasonal median satisfy the assumption that is $F_t$-measurable with no atoms at zero, those can be represented by $S_{d,n}$ and follow the same properties with $S_{d,n}$. The properties are arranged in the following Lemma 1. $BIN(n,p)$ stands for the binomial distribution with the number of trials $n$ and the probability of success $p$.

**Lemma 1.** Let $\{u_t\}$ be a sequence of random variables satisfying Assumption 1. and $\{v_t\}$ be a sequence of $F_t$-measurable random variables with no atoms at zero. Then we can represent (3) as

$S_{d,n} = \sum_{t=1}^{n} sign(u_t) sign(v_{t-d})$

and it satisfies that

(a) $\{S_{d,t}, F_t\}_{t=1}^n$ is a martingale,

(b) $(S_{d,n} + n)/2 \sim BIN(n, 0.5)$.

PROOF.

$$
\begin{aligned}
(a)\ E\left[S_{d,n}|F_{n-1}\right] &= E\left[\sum_{t=1}^{n} sign(u_t)sign(v_{t-d})\middle|F_{n-1}\right] \\
&= E\left[sign(u_n)sign(v_{n-d}) + \sum_{t=1}^{n-1} sign(u_t)sign(v_{t-d})\middle|F_{n-1}\right] \\
&= sign(v_{n-d})E\left[sign(u_n)|F_{n-1}\right] + \sum_{t=1}^{n-1} sign(u_t)sign(v_{t-d}) \\
&= \sum_{t=1}^{n-1} sign(u_t)sign(v_{t-d}) = S_{d,n-1}
\end{aligned}
$$

(b) Let $\{e_t\}$ be iid random variables $t = 1, \ldots, n$, and

$$X_t = sign(e_t) = \begin{cases} 1, & e_t \geq 0,\ p = 0.5 \\ -1, & e_t < 0,\ p = 0.5 \end{cases}.$$

If we let

$$Y_t = \frac{X_t + 1}{2} = \begin{cases} 1 & e_t \geq 0,\ p = 0.5 \\ 0 & e_t < 0,\ p = 0.5 \end{cases},$$

then $Y_t \sim BIN(1, 0.5)$ and

$$\sum_{t=1}^{n} Y_t = \sum_{t=1}^{n} \frac{sign(e_t) + 1}{2} = \frac{\sum_{t=1}^{n} sign(e_t) + n}{2} \sim BIN(n, 0.5).$$

Since $\{S_{d,t}\}_{t=1}^n \sim \{\sum_{k=1}^t sign(e_k)\}_{t=1}^n$, thus $(S_{d,n} + n)/2 \sim BIN(n, 0.5)$. This completes the proof of Lemma 1.

From the Lemma 1, therefore, if

$$S_{d,n} \leq 2B_\alpha(n, 0.5) - n, \tag{4}$$

then we reject the seasonal random walk null hypothesis, and by the central limit theorem we have $S_{d,n}/\sqrt{n} \xrightarrow{\mathcal{D}} N(0,1)$. Thus, (4) is equal to

$$S_{d,n}/\sqrt{n} \leq -z_\alpha, \tag{5}$$

and when it is attained, we reject the null hypothesis. Here, $-z_\alpha$ stands for a lower $\alpha-$th quantile of the standard normal distribution.

## 3 A Monte Carlo Study

In this chapter, we conduct a set of Monte Carlo experiments to investigate the finite sample performance of the seasonal sign tests, $S^s_{d,n}$ and $S^c_{d,n}$, for testing the seasonal unit root null hypothesis $H_0 : \rho = 1$ against the stationary alternative $H_1 : \rho < 1$. Because the DHF test is widely used in practice, we look at its size and power and compare them with those of the seasonal sign tests.

We consider the seasonal AR(1) model with the common mean model

$$\begin{cases} y_t = \mu + u_t \\ u_t = \rho u_{t-d} + e_t, \quad t = 1, \dots, n; \; d = 2, 4, 12, \end{cases} \tag{6}$$

where $y_t$ is the observation at time $t$, $\mu_t$ which is the mean of $y_t$, and the initial conditions $y_t, t = -d+1, -d+2 \dots, 0$, are set to zero. Besides of a normal disturbances we consider a variance mixture(VM) and a Cauchy distributed errors to examine the effects of heavy-tailed errors on both tests. VM has a finite variance, but a Cauchy distribution has an infinite variance.

Table 1: Empirical sizes(%) and size-adjusted powers(%) of the mean-adjusted tests for model $y_t = \rho y_{t-d} + e_t$. The number of replications is 10,000.

| | | | | | Common | | Seasonal | |
|---|---|---|---|---|---|---|---|---|
| $d$ | $n$ | $size^{1)}$ | $e_t$ | $\rho$ | $DHF^c$ | $S^c_{d,n}$ | $DHF^s$ | $S^s_{d,n}$ |
| 4 | 120 | 4.12 | $N(0,1)$ | 1 | 4.99 | 4.12 | 5.07 | 4.10 |
| | | | | 0.99 | 10.17 | 10.21 | 7.71 | 8.76 |
| | | | | 0.95 | 62.11 | 36.72 | 19.86 | 26.81 |
| | | | | 0.90 | 98.78 | 67.18 | 56.92 | 51.06 |
| | | | | 0.80 | 100.00 | 94.85 | 99.42 | 86.25 |
| | | | $VM(1,25)$ | 1 | 4.86 | 4.17 | 6.82 | 4.18 |
| | | | | 0.99 | 9.96 | 12.89 | 7.91 | 10.69 |
| | | | | 0.95 | 61.95 | 57.94 | 17.52 | 38.60 |
| | | | | 0.90 | 98.40 | 87.06 | 53.23 | 68.24 |
| | | | | 0.80 | 99.99 | 99.07 | 99.02 | 94.81 |
| | | | *Cauchy* | 1 | 3.06 | 4.14 | 15.02 | 4.08 |
| | | | | 0.99 | 12.22 | 90.35 | 5.19 | 52.42 |
| | | | | 0.95 | 85.75 | 99.85 | 1.62 | 73.57 |
| | | | | 0.90 | 98.47 | 99.90 | 2.26 | 87.47 |
| | | | | 0.80 | 99.44 | 99.95 | 21.95 | 98.19 |

1) *size* represents the exact size of $S^c_{d,n}$ and $S^s_{d,n}$.

2) DHF$^c$ and DHF$^s$ denote common and seasonal mean adjusted test, respectively.

We use the recursive median adjustment for both test statistics. So and Shin(2001) point out that the recursive mean or median adjustment improves the power of a test. The reader is also referred to So and Shin(1999) and Shin and So(2001). If we let Let $\hat{\mu}_{c,t}$ be the common recursive median of $y_t$, where $y_{(1)} \leq \dots \leq y_{(t)}$ are the order statistics of $y_i, i = 1, \dots, t$. Then $\hat{\mu}_{c,t} = y_{((t+1)/2)}$

for odd $t$ and $\hat{\mu}_{c,t} = y_{(t/2)}$ for even $t$. The $\hat{\mu}_{s,t}$ denotes the seasonal recursive median of $y_t$ and is the median of the observations observed in the same season until the $t$-th observation.

Table 2: Empirical sizes(%) and size-adjusted powers(%) of the mean-adjusted tests for model $y_t = \rho y_{t-d} + e_t$. The number of replications is 10,000.

| | | | | | Common | | Seasonal | |
|---|---|---|---|---|---|---|---|---|
| $d$ | $n$ | $size^{1)}$ | $e_t$ | $\rho$ | $DHF^c$ | $S^c_{d,n}$ | $DHF^s$ | $S^s_{d,n}$ |
| 4 | 120 | 4.12 | *ARCH* | 1 | 9.84 | 4.17 | 9.66 | 4.13 |
| | | | | 0.99 | 5.97 | 6.17 | 5.02 | 6.17 |
| | | | | 0.95 | 15.45 | 18.43 | 7.81 | 13.09 |
| | | | | 0.90 | 38.00 | 35.62 | 13.84 | 21.49 |
| | | | | 0.80 | 80.20 | 59.67 | 35.28 | 38.85 |

1) *size* represents the exact size of $S^c_{d,n}$ and $S^s_{d,n}$.

2) DHF$^c$ and DHF$^s$ denote common and seasonal mean adjusted test, respectively.

We set the nominal level as 5% for every set of experiments, but $S_{d,n}$ is a discrete test of which the nominal level is 4.12%. Thus we adjust the sizes and powers to 4.12% except for the size of $DHF$. We investigate the characteristic of $DHF$ and $S_{d,n}$ with heavy-tailed disturbances in Table 1 and with ARCH disturbances in Table 2. The common characteristic of Table 1 and Table 2 is that in contrast to $DHF$ the size of $S_{d,n}$ quite close to its exact level for all cases. In Table 1 the powers of $S^c_{d,n}$ and $S^s_{d,n}$ are locally higher than those of $DHF^c$ and $DHF^s$, when $\rho$ is close to 1. Therefore $S_{d,n}$ is locally more powerful than $DHF$ under the heavy-tailed disturbances. Moreover Table 2 displays the result under the ARCH errors, the sizes of DHF are 9.84 and 9.66 for common and seasonal case, respectively. And those are much distorted with its nominal level, 5%. However, the size of $S^c_{d,n}$ and $S^s_{d,n}$ meet those nominal level. Thus $S_{d,n}$ are very applicable when the disturbances are heteroscedastic.

## 4 Conclusion

Under the heavy-tailed errors, ARCH, and heteroscedasticity, regardless of the period of the seasonality, the seasonal sign tests are more robust and locally more powerful than the OLSE-based test, DHF. The sign tests follow an exact null distribution which is binomial, regardless of the period of the seasonality. Therefore the tests are very flexible and is useful for unit root tests not only for random walks, as shown in Campbell and Dufour(1995) and So and Shin(2001), but also for the seasonal autoregressive models.

## References

[1] Campbell, B., and Dufour, J. M. (1995), Exact nonparametric orthogonality and random walk tests, *The Review of Economics and Statistics,* 77, 1-16.

[2] Dickey, D. A., Hasza, D. P., and Fuller, W. A. (1984), "Testing for unit roots in seasonal time series," *Journal of the American Statistician Association,* 79, 355-367.

[3] Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B. S. (1990), Seasonal integration and cointegration, *Journal of Econometrics,* 44, 215-238.

[4] Shin, D. W., and So, B. S. (2001), Recursive mean adjustment for unit root tests, *Journal of Time Series Analysis,* 22, 595-612.

[5] So, B. S. (2001) A new estimator for seasonal autoregressive process, *Journal of the Korean Statistical Society,* 30, 31-39.

[6] So, B. S., and Shin, D. W. (1999), Cauchy estimators for autoregressive processes with applications to unit root tests and confidence intervals, *Econometric Theory,* 15, 165-176.

[7] So, B. S., and Shin, D. W. (1999), Recursive mean adjustment in time-series inferences, *Statistics & Probability Letters,* 43, 65-73.

[8] So, B. S., and Shin, D. W. (2001), An invariant sign test for random walks based on recursive median adjustment, *Journal of Econometrics,* 102, 197-229.