Journal of Statistical and Econometric Methods, Vol.11, No.2, 2022, 1-11
ISSN: 2241-0384 (print), 2241-0376 (online)
https://doi.org/10.47260/jsem/1121
Scientific Press International Limited

Comparing Different Permutation Tests with Dickey-Fuller Tests for Unit Root in the Autoregressive Time Series

Jiexiang Li[1]

Abstract

Three permutation tests based T_n (Li et al. (2013)), KS_n and BKR_n (Blum et al.(1961)) for unit root in the AR(1) time series are investigated and compared to Dickey-Fuller tests with white noise from distributions at different levels of skewness (symmetric distributions such as standard normal; slightly skewed distributions such as Chisq (1); highly right skewed distributions such as Weibull (shape=1/3, scale=1); highly left skewed distributions such as negative lognormal ($\mu = 0, \sigma = 2$)) and two moderately skewed F distributions with numerator degree of freedom 1 and denominator degrees of freedom 7 and 4. As expected, Dickey-Fuller tests overperform the permutation tests when white noise is from symmetric distributions or slightly skewed distributions. The permutation tests based on BKR_n perform at least comparable to and most of the time overperform the permutation tests based on KS_n regardless of the levels of skewness of white noise distributions. The permutation tests based on T_n could not compare to Dickey-Fuller tests when white noise is from symmetric distributions and it could not compare to the permutation tests based on BKR_n when white noise is from skewed distributions.

JEL classification numbers: C12, C14, C15.
Keywords: Permutation tests, Autoregressive, Unit root.

---

[1]Department of Mathematics, College of Charleston, U.S.A.

Article Info: Received: May 17, 2022. Revised: June 11, 2022. Published online: June 20, 2022.

1 Introduction

Let Y_1,...,Y_{n+1} be observations from the AR(1) model

$$Y_t = aY_{t-1} + e_t,$$

where $0 < a < 1$ and white noise e_t is a sequence of independent normally distributed random variables with mean 0 and variance σ^2. For large n, maximum likelihood estimator (MLE) of a is normally distributed with mean a and variance $\frac{1-a^2}{n+1}$. Fuller (1976), Dickey and Fuller (1979) constructed test statistics and tables of critical values for tests of

$$H_0 : a = 1 \quad \text{versus} \quad H_A : 0 < a < 1,$$

which are often referred to as tests for unit root. The hypothesis that $a = 1$ is of interest in applications because it corresponds to the hypothesis that it is appropriate to transform the times series by differencing. For literature on autoregressive processes and tests for unit root, the reader is referred to Brockwell and Davis (1996), Fuller (1976), Dickey and Fuller (1979). In this paper, we will extend Fuller (1976), Dickey and Fuller (1979) so that white noise from the AR(1) model in tests for unit root is not limited to normal distributions. Define $X_t = Y_{t+1} - Y_t$, $t = 1, 2, ..., n$. Then

$$X_t = (a - 1)Y_t + e_{t+1}.$$

Under H_0, $X_t = e_{t+1}$, $t = 1, 2, ..., n$ are independent continuous random variables. Given $n+1$ observations $Y_1, Y_2, ..., Y_{n+1}$, testing for unit root is equivalent to testing X_1,X_2,...,X_n are independent.

Lemma 1.1 *(Li et al. (2013)) Under H_A, for any integer $m \geq 1$,*

$$CORR(X_1, X_{1+m}) = \frac{2a^m - a^{m-1} - a^{m+1}}{2(1-a)} = -\frac{1-a}{2}a^{m-1}.$$

As defined in Li (2013,) $\mathrm{T}_n = \sum_{i=1}^{n-1} X_iX_{i+1}$.

Lemma 1.2 *(Li et al. (2013)) Under H_A, $\frac{T_n}{n-1}$ converges in probability to $EX_1X_2 = -\frac{\sigma^2(1-a)}{1+a}$.*

Under H_0, $\frac{\mathrm{T}_n}{n-1}$ converges in probability to $\mathrm{E}X_1X_2 = 0$, which along with Lemma 1.2 form a basis for testing for unit root.

2 Methods

Define $Z_t = (X_t, X_{t+1})$, $t = 1, 2, ..., n-1$. Assume that Z_t has a continuous joint cumulative distribution function $F(\cdot,\cdot)$ and a continuous marginal cumulative distribution function $F_1(\cdot)$. Observing Lemma 1.1, it is sufficient to test for unit root by testing

$$H_0 : S(\mathbf{x}) = 0 \text{ for all } \mathbf{x} = (x_1, x_2) \in R^2 \tag{2.1}$$

where

$$S(\mathbf{x}) = F(\mathbf{x}) - F_1(x_1)F_1(x_2)$$

versus

$$H_A : S(\mathbf{x}) \neq 0 \text{ for some } \mathbf{x} = (x_1, x_2) \in R^2 \tag{2.2}$$

Clearly,

$$S(\mathbf{x}) = E\{\Pi_{j=1}^2 I(X_j \leq x_j)\} - \Pi_{j=1}^2 E\{I(X_j \leq x_j)\}$$

where I(A) is the indicator function of the event A. Define,

$$S_n(\mathbf{x}) = (n-1)^{-1} \sum_{t=1}^{n-1} \Pi_{j=1}^2 I(X_{t+j-1} \leq x_j) - \Pi_{j=1}^2 \{(n-1)^{-1} \sum_{t=1}^{n-1} I(X_{t+j-1} \leq x_j)\}$$

for any $\mathbf{x} = (x_1, x_2) \in R^2$. Consider Kolmogorov-Smirnov statistic

$$\mathrm{KS}_n = \max_{1 \leq i \leq n-1} |S_n(x_i, x_{i+1})|$$

and

$$\mathrm{BKR}_n = \frac{\sum_{i=1}^{n-1} S_n^2(x_i, x_{i+1})}{n-1}$$

given $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$. Since KS_n and BKR_n takes small values under H_0 and larger values under H_A, it forms a basis for testing for unit root. For reference, Skaug and Tjϕstheim (1993) and Blum et al.(1961) investigated nonparametric tests of serial independence based on the fact that the null hypothesis of independence holds if and only if the joint distribution function equals the product of the marginal distribution functions.

3 Main Results

3.1 Steps Used in Permutation Tests

Throughout the paper, we assume that white noise is a sequence of independent identically distributed continuous random variables with mean zero and finite variance σ^2.

Permutation tests are carried out as that is described in Li et al. (2013). For observations $X_1, X_2, ..., X_n$, there are a total of $n!$ permutations. This test is limited by prohibitive calculations and takes a large amount of time to execute if n is a large number. Instead of using all $n!$ permutations to compute the $p-$ value , we obtain a random sample of R permutations. The statistics computed from each permuted sequence $X_{1l}, ..., X_{(n)l}$ are referred to as $\mathrm{T}_{n,l}$, $\mathrm{KS}_{n,l}$ and $\mathrm{BKR}_{n,l}$, and the statistics computed from the observations are referred to as $\mathrm{T}_{n,obs}$, $\mathrm{KS}_{n,obs}$ and $\mathrm{BKR}_{n,obs}$. Note that under H_0, $\mathrm{T}_{n,l}$, $\mathrm{KS}_{n,l}$ and $\mathrm{BKR}_{n,l}$, $1 \leq l \leq R$, are equally likely. The steps used in permutation tests are outlined below.

1. Set a predetermined level α. Compute test statistics $\mathrm{T}_{n,l}$, $\mathrm{KS}_{n,l}$ and $\mathrm{BKR}_{n,l}$ for each sampled permutation $1 \leq l \leq R$ and observed test statistics $\mathrm{T}_{n,obs}$, $\mathrm{KS}_{n,obs}$ and $\mathrm{BKR}_{n,obs}$ based on original (not permuted) sample.

2. Compute $p-$value based on T_n as the proportion of $\mathrm{T}_{n,l}$'s less than or equal to $\mathrm{T}_{n,obs}$; $p-$value based on KS_n as proportion of $\mathrm{KS}_{n,l}$'s greater than or equal to $\mathrm{KS}_{n,obs}$; $p-$value based on BKR_n as the proportion of $\mathrm{BKR}_{n,l}$'s greater than or equal to $\mathrm{BKR}_{n,obs}$, that is,

$$p - \text{value based on } \mathrm{T}_n = \frac{\sum_{l=1}^{R} I(\mathrm{T}_{n,l} \leq \mathrm{T}_{n,obs})}{R};$$

$$p - \text{value based on } \mathrm{KS}_n = \frac{\sum_{l=1}^{R} I(\mathrm{KS}_{n,l} \geq \mathrm{KS}_{n,obs})}{R};$$

$$p - \text{value based on } \mathrm{BKR}_n = \frac{\sum_{l=1}^{R} I(\mathrm{BKR}_{n,l} \geq \mathrm{BKR}_{n,obs})}{R}.$$

Conclude that the tests are statistically significant if the corresponding $p-$values are less than or equal to α.

3.2 Consistency of Permutation Tests

Consistency of a hypothesis test is a desirable property. In this section, we will show that the permutation tests based on random sampling of R permutations are consistent. In Li et al. (2013), the permutation test based on T_n was shown to be consistent. We will focus on the proof of consistency of the hypothesis test based on KS_n because the proof of consistency of the hypothesis test based on BKR_n follows along.

Theorem 3.1 *Suppose H_a is an arbitrary simple hypothesis that the autoregressive parameter a is between 0 and 1, that is $H_a \in H_A$. Then for permutation tests based on statistics KS_n and BKR_n defined above,*

$$P_{H_a}[\textit{Reject } H_0] \to 1$$

as $n \to \infty$.

For large n, the permuted sequence $(X_{1l}, ..., X_{nl})$, $1 \leq l \leq R$, "behaves" like a sequence of independent random variables under H_a (Li et al. (2013)). Hence, we have

Lemma 3.1 *Under H_a, $KS_{n,l} \to 0$ and $BKR_{n,l} \to 0$ a.s. as $n \to \infty$ for all $1 \leq l \leq R$.*

Under H_a, $\mathrm{KS}_{n,obs} \to \sup_{(x_1,x_2)\in R^2} |S(x_1, x_2)| := \beta > 0$ a.s. as $n \to \infty$ following Newman (1984) and Jabbari et al. (2009) and Lemma 1.1. Therefore,

$$P_{H_a}(\mathrm{KS}_{n,obs} > \frac{\beta}{2}) \to 1 \tag{3.1}$$

Note that under H_a, $\mathrm{BKR}_{n,obs} \to \mathrm{E}S^2(X_1, X_2) := \alpha > 0$ a.s. as $n \to \infty$. Therefore,

$$P_{H_a}(\mathrm{BKR}_{n,obs} > \frac{\alpha}{2}) \to 1$$

If $\mathrm{KS}_{n,obs} > \frac{\beta}{2} > \mathrm{KS}_{n,l}$ for all $1 \leq l \leq R$, which means the fraction of $\mathrm{KS}_{n,l}$'s that are greater than or equal to $\mathrm{KS}_{n,obs}$ is zero. Consequently, the p-value is zero and H_0 is rejected. Therefore, we have

$$\{\mathrm{KS}_{n,obs} > \frac{\beta}{2}\} \cap_{l=1}^{R} \{\mathrm{KS}_{n,l} < \frac{\beta}{2}\} \subset \text{Reject } H_0. \tag{3.2}$$

Note for any two sequences of events A_n and B_n, $P(\overline{A_n \cap B_n}) = P(\bar{A}_n \cup \bar{B}_n) \leq P(\bar{A}_n) + P(\bar{B}_n)$. Therefore, if $P(A_n) \to 1$ and $P(B_n) \to 1$, we have $P(A_n \cap B_n) \to 1$. Hence, in hypothesis test based on KS_n, $P_{H_a}(\text{Reject } H_0) \to 1$ based on Lemma 3.1, (3.1) and (3.2).

4 Simulation Studies

We generate $n+1$ observations from the real valued $AR(1)$ model

$$Y_t = aY_{t-1} + e_t,$$

We focus on six white noise distributions: (1) standard normal; (2) χ^2 with 1 degree of freedom; (3) Weibull with scale parameter=1 and shape parameter=$\frac{1}{3}$; (4) negative lognormal with $\mu = 0$ and $\sigma = 2$; (5) F distribution with numerator degree of freedom=1 and denominator degrees of freedom=7; (6) F distribution with numerator degree of freedom=1 and denominator degrees of freedom=4. Note that mean of χ^2 with df= 1 is 1; mean of Weibull with scale parameter=1 and shape parameter=$\frac{1}{3}$ is 3!=6; mean of negative lognormal with $\mu = 0$ and $\sigma = 2$ is $-e^2$; mean of F (1,7) is $\frac{7}{(7-2)} = \frac{7}{5}$ and mean of F (1,4) is $\frac{4}{(4-2)} = 2$. We will shift distributions (2), (3), (4), (5) and (6) by subtracting their corresponding means so that the means of distributions (2), (3), (4), (5) and (6) after shifting are equal to 0. Note also that shifted distributions (2), (3), (4), (5) and (6) have finite variance. We will generate white noise from distribution(1) and shifted distributions (2), (3), (4), (5) and (6). The six white noise distributions are presented in Figure 1. In our simulations, we randomly select 100 permutations and repeat each test 10,000 for powers and probabilities of type I error. For Dickey-Fuller tests, the test statistics are obtained by standardizing, under H0: $a = 1$, the least square estimator $\hat{a}$ from simple linear regression while regressing Y_t on Y_{t-1} without the constant term. We use critical values -1.947 for sample size 50, -1.944 for sample size 100 and -1.942 for sample size 250 in Dickey-Fuller tests. In our simulations, we consider $a = 0.8$, 0.9, 0.95, 0.99, 1 and $n = 50$, 100, 250. We choose nominal level of significance $\alpha = 0.05$. We summarize the proportions of rejecting the null hypothesis out of 10,000 simulations based on the three permutation test statistics T_n, KS_n and BKR_n and Dickey-Fuller tests, which are our estimated powers ($a \neq 1$) and estimated probabilities of type I error ($a = 1$). Based on our simulations, we recommend Dickey-Fuller tests for unit root when white noise is from symmetric or slightly skewed distributions (Table 1). We recommend permutation tests based on BKR_n for unit root when white noise is from heavily skewed distributions (Table 2). For white noise distributions with moderate skewness, performances of Dickey-Fuller tests and permutation tests based on BKR_n depend on the value of a and sample size n (Table 3). For white noise from shifted F (1,7), in some cases ($a = 0.8$ and $n = 50, 100$; $a = 0.9$ and $n = 100, 250$; $a = 0.95$ and $n = 250$), Dickey-Fuller tests are more powerful than permutations tests based on BKR_n; in other cases, permutation tests based on BKR_n are slightly more powerful than or comparable to Dickey-Fuller tests. For white noise from shifted F (1,4), permutations tests based on BKR_n are more powerful than or at least comparable to Dickey-Fuller tests except $a = 0.8$ and $n = 50, 100$.

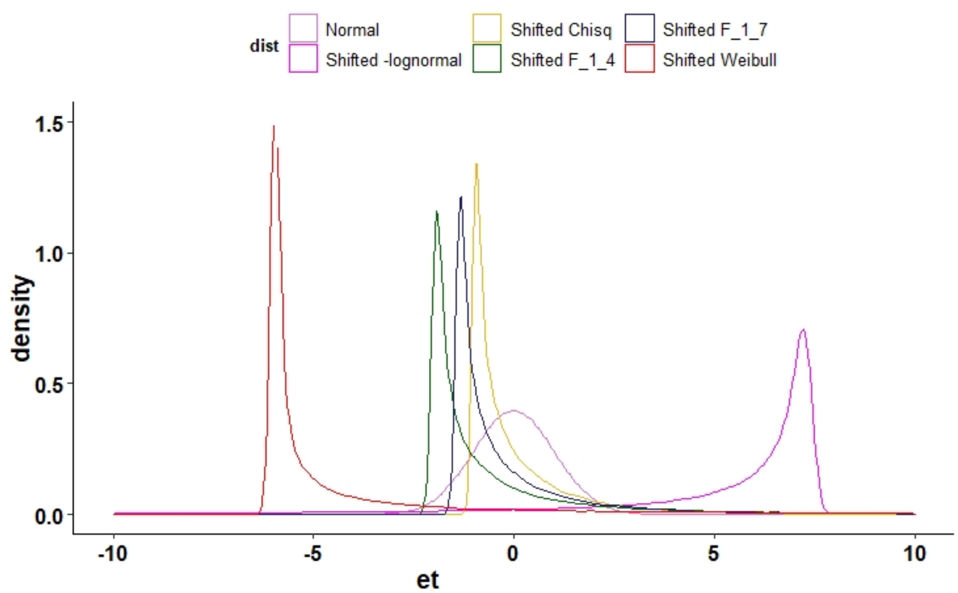

Figure 1: white noise distributions under investigation

Table 1: Simulated power for white noise from Standard Normal or χ^2 distribution

		standard normal				**shifted χ^2**			
Sample	a	T_n	KS_n	BKR_n	Dickey-Fuller	T_n	KS_n	BKR_n	Dickey-Fuller
50	**0.8**	0.15	0.07	0.07	0.78	0.2	0.19	0.27	0.77
	0.9	0.09	0.06	0.05	0.32	0.11	0.13	0.17	0.31
	0.95	0.07	0.06	0.06	0.15	0.07	0.09	0.11	0.14
	0.99	0.06	0.06	0.07	0.07	0.05	0.06	0.07	0.06
	1	0.06	0.06	0.06	0.05	0.06	0.06	0.06	0.05
100	**0.8**	0.25	0.1	0.11	1	0.31	0.36	0.52	0.99
	0.9	0.12	0.06	0.06	0.77	0.15	0.22	0.29	0.76
	0.95	0.08	0.06	0.06	0.32	0.09	0.13	0.16	0.31
	0.99	0.07	0.06	0.06	0.08	0.06	0.06	0.07	0.07
	1	0.06	0.05	0.05	0.05	0.06	0.06	0.06	0.05
250	**0.8**	0.49	0.21	0.27	1	0.55	0.83	0.96	1
	0.9	0.2	0.08	0.09	1	0.24	0.51	0.67	1
	0.95	0.11	0.06	0.06	0.9	0.12	0.28	0.34	0.9
	0.99	0.07	0.06	0.06	0.15	0.07	0.07	0.09	0.15
	1	0.06	0.06	0.06	0.05	0.06	0.06	0.06	0.05

Table 2: Simulated power for white noise from Weibull or -lognormal distribution

		Shifted Weibull				shifted negative lognormal			
Sample	a	T_n	KS_n	BKR_n	Dickey-Fuller	T_n	KS_n	BKR_n	Dickey-Fuller
50	**0.8**	0.25	0.92	0.96	0.66	0.23	0.76	0.84	0.6
	0.9	0.12	0.89	0.94	0.22	0.11	0.72	0.76	0.19
	0.95	0.07	0.82	0.86	0.07	0.07	0.6	0.64	0.06
	0.99	0.06	0.45	0.48	0.03	0.06	0.19	0.22	0.02
	1	0.06	0.06	0.06	0.02	0.06	0.05	0.06	0.02
100	**0.8**	0.52	1	1	0.94	0.52	0.97	0.99	0.91
	0.9	0.25	1	1	0.7	0.24	0.95	0.97	0.63
	0.95	0.13	0.98	0.99	0.24	0.11	0.91	0.93	0.2
	0.99	0.07	0.79	0.81	0.04	0.06	0.47	0.49	0.04
	1	0.06	0.06	0.06	0.03	0.06	0.06	0.06	0.02
250	**0.8**	0.8	1	1	1	0.82	1	1	1
	0.9	0.5	1	1	0.99	0.52	1	1	0.98
	0.95	0.25	1	1	0.84	0.26	1	1	0.79
	0.99	0.08	1	1	0.1	0.08	0.93	0.93	0.08
	1	0.06	0.06	0.06	0.03	0.06	0.06	0.06	0.03

Table 3: Simulated power for white noise from two F-distributions

		Shifted F (1,7)				Shifted F (1,4)			
Sample	a	T_n	KS_n	BKR_n	Dickey-Fuller	T_n	KS_n	BKR_n	Dickey-Fuller
50	**0.8**	0.22	0.34	0.44	0.74	0.24	0.52	0.63	0.69
	0.9	0.12	0.26	0.31	0.28	0.12	0.42	0.49	0.23
	0.95	0.08	0.17	0.2	0.11	0.08	0.28	0.32	0.09
	0.99	0.06	0.06	0.08	0.05	0.06	0.09	0.11	0.03
	1	0.06	0.06	0.06	0.04	0.06	0.06	0.06	0.02
100	**0.8**	0.38	0.62	0.79	0.99	0.44	0.81	0.91	0.97
	0.9	0.17	0.47	0.56	0.75	0.2	0.71	0.78	0.7
	0.95	0.11	0.29	0.34	0.29	0.11	0.53	0.59	0.24
	0.99	0.06	0.08	0.1	0.07	0.06	0.15	0.17	0.05
	1	0.06	0.06	0.06	0.04	0.06	0.05	0.06	0.03
250	**0.8**	0.64	0.98	1	1	0.73	1	1	1
	0.9	0.3	0.87	0.94	1	0.41	0.98	0.99	0.99
	0.95	0.15	0.67	0.74	0.89	0.21	0.92	0.95	0.84
	0.99	0.07	0.15	0.17	0.13	0.07	0.4	0.42	0.1
	1	0.06	0.06	0.06	0.04	0.06	0.06	0.06	0.03

References

[1] Blum, J. R., Kiefer, J. and Rosenblatt, M. (1961), Distribution free tests of independence based on the sample distribution function, *The Annals of Mathematical Statistics*, **32**(2), 485-498. 405-417.

[2] Brockwell, P. J. and Davis, R. A. (1996), *Introduction to time series and forecasting*, Springer.

[3] Fuller, Wayne A. (1976), *Introduction to statistical times series*, New York: John Wiley & Sons.

[4] Dickey, David A. and Fuller, Wayne A. (1979), Distributions of estimates for Autoregressive Time series with a unit root, *Journal of American Statistical Association*, **74**(365), 427-431.

[5] Jabbari H., Azarnoosh, H. A. and Fakoor V. (2009), Almost sure convergence of two-dimensional distribution function under negative association, *Journal of Applied Probability and Statistics*, **4**(2), 157-166.

[6] Li, J., Tran, L. and Niwitpong, S. (2013), A permutation test for unit root in an autoregressive model, *Applied Mathematics*, 4, 1629-1634.

[7] Newman, C. M. (1984), Asymptotic independence and limit theorems for positively and negatively dependent random variables, *Inequalities in statistics and probability*, IMS Lecture Notes-Monograph Series, 5, 127-140.

[8] Skaug, H. J. and Tjϕstheim(1993), A nonparamentric test of serial independence based on the empirical distribution function, *Biometrika*, **80**(3), 591-602.