Iran J Sci Technol Trans Sci (2017) 41:665–669
DOI 10.1007/s40995-017-0288-8

RESEARCH PAPER

Testing the Difference between Two Independent Time Series Models

Mohammad Reza Mahmoudi[1] **· Mohsen Maleki**[2] **· Abbas Pak**[3]

Received: 13 August 2015 / Accepted: 17 February 2016 / Published online: 19 September 2017

Abstract In some situations, for example in biology, economic, electronic, finance and management, researchers wish to determine whether the two time series are generated by the same stochastic mechanism or their random behavior differs. In this work, the asymptotic distribution for the difference of two independent ARMA coefficients is established. The presented method can be used to derive the asymptotic confidence set for the difference of coefficients and hypothesis testing for the equality of two time series. Then the Monte Carlo simulation study is provided to investigate the performance of proposed method. The performance of the new method is comparable with alternative method.

Keywords Asymptotic · ARMA processes · Simulation · Simultaneous inference · Time series

✉ Mohammad Reza Mahmoudi
mahmoudi.m.r@fasau.ac.ir

Mohsen Maleki
stud.stat@yahoo.com

Abbas Pak
rezastat2211391@gmail.com

1 Department of Statistics, Faculty of Science, Fasa University, Fasa, Iran

2 Shiraz University, Shiraz, Iran

3 Department of Computer Sciences, Faculty of Mathematical Sciences, Shahrekord University, P. O. Box 115, Shahrekord, Iran

1 Introduction

In some situations, for example in biology, economic, electronic, finance and management researches, we wish to determine whether the two time series are generated by the same stochastic mechanism or their random behavior differs. The comparison of two time series models has been studied in both time- and frequency-domain methods. Coates and Diggle (1986), Diggle and Fisher (1991), Dargahi-Noubary (1992), Diggle and Al Wasel (1997), Kakizawa et al. (1998), Maharaj (2002) and Caiado et al. (2006) studied the comparison and discrimination of time series with equal length using spectral analysis approaches. Classification and clustering analysis of time series with unequal length have done by Camacho et al. (2004). They truncated data and used information about truncated time series spectra to compare them. Caiado et al. (2006) proposed a method based on periodogram to compare two unequal length time series. They calculated the periodogram at different Fourier frequencies and proposed nonparametric and parametric test statistics to test the hypothesis that the two series with different lengths are generated by the same stochastic mechanism.

In this work, the asymptotic distribution for the difference of two ARMA coefficients is presented. This can be applied to construct confidence bands, and a test to compare two time series models. Also, we present the results of Monte Carlo simulation study on the performance of the proposed method. Finally, we compare this method with approach which was introduced by Caiado et al. (2006).

2 Large Samples Inference

For two independent populations, suppose $\left\{X_t^{(1)}, t=1,\ldots,n_1\right\}$ and $\left\{X_t^{(2)}, t=1,\ldots,n_2\right\}$ are causal invertible ARMA(p,q) processes,

$$X_t^{(i)} - \phi_1^{(i)} X_{t-1}^{(i)} - \cdots - \phi_p^{(i)} X_{t-p}^{(i)} = Z_t^{(i)} + \theta_1^{(i)} Z_{t-1}^{(i)} + \cdots + \theta_q^{(i)} Z_{t-q}^{(i)}, \left\{Z_t^{(i)}\right\} \sim \text{IID}(0, \sigma^2), \quad i = 1, 2. \tag{2.1}$$

Assume $\hat{\boldsymbol{\beta}}_i = (\hat{\phi}_1^{(i)}, \ldots, \hat{\phi}_p^{(i)}, \hat{\theta}_1^{(i)}, \ldots, \hat{\theta}_q^{(i)})' = (\hat{\phi}^{(i)}, \hat{\theta}^{(i)})'$, $i = 1, 2$, and $\hat{\sigma}^2$ are the maximum likelihood estimators of $\boldsymbol{\beta}_i = (\phi^{(i)}, \boldsymbol{\theta}^{(i)})', i = 1, 2$, and σ^2, respectively.

We are interested in testing $H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ versus $H_1 : \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$. If H_0 is rejected, then the two time series models are different, and if H_0 is not rejected, then there is no significant difference between two models and two time series have same stochastic mechanism.

Lemma 2.1 *Under the assumptions of model* (2.1), *for* $i = 1, 2$,

$$\sqrt{n}\left(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i\right) \to N_{p+q}(0, \text{V}_i) \quad \text{as } n \to \infty,$$

where $n = \min(n_1, n_2)$, *and the asymptotic covariance matrix* V_i *can be computed from*

$$\text{V}_i = \sigma^2 \begin{bmatrix} E\left(U_t^{(i)} U_t^{'(i)}\right) & E\left(U_t^{(i)} V_t^{'(i)}\right) \\ E\left(V_t^{(i)} U_t^{'(i)}\right) & E\left(V_t^{(i)} V_t^{'(i)}\right) \end{bmatrix}^{-1},$$

where $\boldsymbol{U}_t^{(i)} = (U_t^{(i)}, \ldots, U_{t+1-p}^{(i)})'$, $\boldsymbol{V}_t^{(i)} = (V_t^{(i)}, \ldots, V_{t+1-q}^{(i)})'$ *and* $\{U_t\}$, $\{V_t\}$ *are the autoregressive processes,*

$$\phi(B) U_t = Z_t,$$

and

$$\theta(B) V_t = Z_t.$$

Proof The outline of the proof is given in Brockwell and Davis (1991). □

We are interested in making inference about the parameter $\boldsymbol{\beta} = \boldsymbol{\beta}_1 - \boldsymbol{\beta}_2$. We will apply a methodology similar to the works of Mahmoudi and Mahmoudi (2013a, b), and Mahmoudi et al. (2014). Since $\hat{\boldsymbol{\beta}}_i, i = 1, 2$, are estimators for $\boldsymbol{\beta}_i, i = 1, 2$, $\boldsymbol{b} = \hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2$ is a reasonable estimator for the parameter $\boldsymbol{\beta}$.

Now, by using the Cramer's theorem and Slutsky's theorem, the asymptotic distribution for $\boldsymbol{b}$ is therefore given by

$$\sqrt{n}(\boldsymbol{b} - \boldsymbol{\beta}) \to N_{p+q}(0, \Sigma), \quad \text{as } n \to \infty,$$

where

$$\Sigma = \text{V}_1 + \text{V}_2.$$

Consequently we have

$$T_n = n(\boldsymbol{b} - \boldsymbol{\beta})' \Sigma^{-1} (\boldsymbol{b} - \boldsymbol{\beta}) \to \chi^2(p+q), \quad \text{as } n \to \infty, \tag{2.2}$$

where $\chi^2(p+q)$ is Chi-square distribution with $(p+q)$ degrees of freedom. This result can be used to construct asymptotic confidence bands and hypothesis testing.

2.1 Asymptotic Confidence Bands

Note that the parameter Σ in T_n given by (2.2) depends on the unknown parameters V_1, V_2 and σ^2, so it cannot be used as a pivotal quantity for the parameter $\boldsymbol{\beta}$. We use common estimators of above parameters; therefore, we have

$$T_n^* = n(\boldsymbol{b} - \boldsymbol{\beta})' \hat{\Sigma}^{-1} (\boldsymbol{b} - \boldsymbol{\beta}) \to \chi^2(p+q) \quad \text{as } n \to \infty, \tag{2.3}$$

where $\hat{\Sigma} = \hat{\text{V}}_1 + \hat{\text{V}}_2$. (See Ferguson 1996.)

Now, T_n^* can be used as a pivotal quantity to construct asymptotic confidence bands for $\boldsymbol{\beta}$,

$$P\left\{\boldsymbol{\beta} : n(\boldsymbol{b} - \boldsymbol{\beta})' \hat{\Sigma}^{-1} (\boldsymbol{b} - \boldsymbol{\beta}) \leq \chi^2_{1-\alpha}(p+q)\right\} = 1 - \alpha, \tag{2.4}$$

where $\chi^2_{1-\alpha}(p+q)$ is $(1-\alpha)$ percentile of $\chi^2(p+q)$ distribution.

2.2 Hypothesis Testing

As mentioned, hypothesis testing about $\boldsymbol{\beta}$ is important in practice. For instance, the assumption $\boldsymbol{\beta} = 0$ is equivalent to the assumption $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ (equality of two models). In general, to test $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$, the test statistic can be

$$T_0 = n(\boldsymbol{b} - \boldsymbol{\beta}_0)' \hat{\Sigma}^{-1} (\boldsymbol{b} - \boldsymbol{\beta}_0). \tag{2.5}$$

Under null hypothesis, T_0 has asymptotically a $\chi^2(p+q)$ distribution. Therefore, the critical region for a test of size α is $T_0 > \chi^2_{1-\alpha}(p+q)$.

Remark In application, the orders of model for each time series are selected by Akaike information-corrected criterion (AICC) or Bayesian information criterion (BIC) selection criterions. If ARMA(p_1, q_1) and ARMA(p_2, q_2) are best models to series 1 and 2, respectively, then we assume $p = \max\{p_1, p_2\}$, $q = \max\{q_1, q_2\}$ and fit ARMA(p, q) model to each time series.

3 Simulation Study

In this section, numerous data sets are generated and analyzed to investigate the performance of proposed method. Estimated type I error probabilities $(\hat{\alpha})$ and empirical powers $(\hat{\pi})$ are computed for different values of

$$(n1, n2) = \{(100, 50), (150, 75), (200, 100), (500, 250), (1000, 500), (2000, 1000)\}$$

based on 1000 replications. Also we compare the performance of proposed method with the method which was presented by Caiado et al. (2006). The methodology for computing the test statistic based on their method runs as follows:

(i) Compute the periodogram ordinates of $\left\{X_t^{(k)}\right\}$ by

$$P_{X_t^{(k)}}(\omega_j) = (2\pi n_k)^{-1}\left|\sum_{t=1}^{n_k} X_t^{(k)} e^{-it\omega_j}\right|^2,$$

where $\omega_j = 2\pi j/n_k$, for $j = 1, \ldots, m_k$, with $m_k = [n_k/2]$, the largest integer less than or equal to $n_k/2$, and the frequency ω_j is in the range $[-\pi, \pi]$.
Without loss of generality, let $r = [pm_1/m_2]$ be the largest integer less than or equal to pm_1/m_2 for $p = 1, \ldots, m_2$, and $m_2 < m_1$.

(ii) Normalize the periodograms, dividing by the sample variances and let

$$\mathrm{NP}_{X_t^{(k)}}(\omega_p) = \frac{P_{X_t^{(k)}}(\omega_j)}{\widehat{\mathrm{Var}}(X_t^{(k)})}.$$

(iii) Define the sample mean and sample variance of the log-normalized periodogram by

$$\bar{X}_{\mathrm{LNP},k} = \frac{1}{m}\sum_{p=1}^{m} \log\left(\mathrm{NP}_{X_t^{(k)}}(\omega_p)\right),$$

and

$$S^2_{\mathrm{LNP},k} = \frac{1}{m}\sum_{p=1}^{m}\left(\log\left(\mathrm{NP}_{X_t^{(k)}}(\omega_p)\right) - \bar{X}_{\mathrm{LNP},k}\right)^2,$$

where m is the number of periodogram ordinates of the series with shorter length (in this case $m = m_2$).

(iv) Compute the test statistic

$$D_{\mathrm{NP}} = \frac{\frac{1}{m}\sum_{p=1}^{m}\left(\log\left(\mathrm{NP}_{X_t^{(1)}}(\omega_p)\right) - \log\left(\mathrm{NP}_{X_t^{(2)}}(\omega_p)\right)\right)}{\sqrt{\left(S^2_{\mathrm{LNP},1} + S^2_{\mathrm{LNP},2}\right)/m}},$$

for comparison of the log-normalized periodograms of the two series.

D_{NP} has asymptotically a standard normal distribution. Therefore, the critical region for a test of level α is $|D_{\mathrm{NP}}| > Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is $(1-\alpha/2)$ percentile of standard normal distribution.

Example 1 (AR(1)). Consider the AR(1) model

$$X_t^{(i)} - \phi_1^{(i)} X_{t-1}^{(i)} = Z_t^{(i)}, \quad \left\{Z_t^{(i)}\right\} \sim \mathrm{IID}(0, 1), \quad i = 1, 2.$$

Therefore, we have

$$\sqrt{n}(b - \beta) \to N_1(0, \Sigma), \quad \text{as } n \to \infty,$$

where $\Sigma = \mathrm{V}_1 + \mathrm{V}_2$ and $\mathrm{V}_i = \left(1 - \left(\phi_1^{(i)}\right)^2\right)$.

We consider $\phi_1^{(1)} = 0.5$ for the first population. For the second population, we consider $\phi_1^{(2)} = 0.1, 0.3, 0.5, 0.7$ and 0.9.

The two rows opposite of $\phi_1^{(2)} = 0.5$ in Table 1 report the values of $\hat{\alpha}$ for proposed method (above row) and D_{NP} test statistic (bottom row), respectively. Other rows give the values of $\hat{\pi}$ for the other values of $\phi_1^{(2)}$.

Example 2 (MA(1)). Consider the MA(1) model

$$X_t^{(i)} = Z_t^{(i)} + \theta_1^{(i)} Z_{t-1}^{(i)}, \quad \left\{Z_t^{(i)}\right\} \sim \mathrm{IID}(0, 1), \quad i = 1, 2.$$

Therefore, we have

$$\sqrt{n}(b - \beta) \to N_1(0, \Sigma), \quad \text{as } n \to \infty,$$

where $\Sigma = \mathrm{V}_1 + \mathrm{V}_2$ and $\mathrm{V}_i = \left(1 - \left(\theta_1^{(i)}\right)^2\right)$.

We consider $\theta_1^{(1)} = 0.5$ for the first population. For the second population, we consider $\theta_1^{(2)} = 0.1, 0.3, 0.5, 0.7$ and 0.9.

The two rows opposite of $\theta_1^{(2)} = 0.5$ in Table 2 report the values of $\hat{\alpha}$ for proposed method (above row) and D_{NP} test statistic (bottom row), respectively. Other rows give the values of $\hat{\pi}$ for the other values of $\theta_1^{(2)}$.

Example 3 (ARMA(1,1)). Consider the ARMA(1,1) model

$$X_t^{(i)} - \phi_1^{(i)} X_{t-1}^{(i)} = Z_t^{(i)} + \theta_1^{(i)} Z_{t-1}^{(i)}, \quad \left\{Z_t^{(i)}\right\} \sim \mathrm{IID}(0, 1), \quad i = 1, 2.$$

Therefore, we have

$$\sqrt{n}(\boldsymbol{b} - \boldsymbol{\beta}) \to N_2(0, \boldsymbol{\Sigma}), \quad \text{as } n \to \infty,$$

Table 1 Estimates of power and size of 0.05 test of significance for AR(1), $\phi_1 = 0.5$ versus AR(1), $\phi_1 = 0.1, 0.3, 0.5, 0.7$ and 0.9.

$\phi_1^{(2)}$	Method	(n_1, n_2)					
		(100, 50)	(150, 75)	(200, 100)	(500, 250)	(1000, 500)	(2000, 1000)
0.1	Proposed	0.784	0.898	0.999	1.000	1.000	1.000
	D_{NP}	0.178	0.227	0.304	0.718	0.934	1.000
0.3	Proposed	0.486	0.643	0.785	0.964	0.998	1.000
	D_{NP}	0.118	0.178	0.167	0.378	0.632	0.876
0.5	Proposed	0.068	0.064	0.059	0.056	0.052	0.052
	D_{NP}	0.074	0.071	0.068	0.059	0.058	0.055
0.7	Proposed	0.464	0.656	0.812	0.996	1.000	1.000
	D_{NP}	0.234	0.285	0.387	0.785	0.957	1.000
0.9	Proposed	0.814	0.943	1.000	1.000	1.000	1.000
	D_{NP}	0.718	0.854	0.954	0.994	1.000	1.000

Table 2 Estimates of power and size of 0.05 test of significance for MA(1), $\theta_1 = 0.5$ versus MA(1), $\theta_1 = 0.1, 0.3, 0.5, 0.7$ and 0.9.

$\theta_1^{(2)}$	Method	(n_1, n_2)					
		(100, 50)	(150, 75)	(200, 100)	(500, 250)	(1000, 500)	(2000, 1000)
0.1	Proposed	0.678	0.823	0.966	0.999	1.000	1.000
	D_{NP}	0.186	0.216	0.326	0.698	0.898	0.966
0.3	Proposed	0.423	0.589	0.765	0.896	0.934	0.986
	D_{NP}	0.143	0.187	0.198	0.390	0.628	0.899
0.5	Proposed	0.066	0.058	0.052	0.050	0.049	0.049
	D_{NP}	0.073	0.068	0.059	0.058	0.058	0.057
0.7	Proposed	0.432	0.634	0.796	0.923	1.000	1.000
	D_{NP}	0.222	0.295	0.402	0.800	0.949	0.996
0.9	Proposed	0.823	0.953	1.000	1.000	1.000	1.000
	D_{NP}	0.743	0.798	0.906	0.986	0.994	1.000

Table 3 Estimates of power and size of 0.05 test of significance for ARMA(1,1), $\phi_1 = 0.2$, $\theta_1 = -0.5$ versus ARMA(1,1), $\phi_1 = 0.2, \theta_1 = -0.1, -0.3, -0.5, -0.7$ and -0.9.

$\theta_1^{(2)}$	Method	(n_1, n_2)					
		(100, 50)	(150, 75)	(200, 100)	(500, 250)	(1000, 500)	(2000, 1000)
−0.1	Proposed	0.731	0.904	0.996	1.000	1.000	1.000
	D_{NP}	0.205	0.294	0.364	0.703	0.934	0.988
−0.3	Proposed	0.521	0.643	0.722	0.828	1.000	1.000
	D_{NP}	0.112	0.134	0.138	0.304	0.498	0.862
−0.5	Proposed	0.055	0.053	0.054	0.052	0.051	0.050
	D_{NP}	0.054	0.053	0.053	0.052	0.048	0.052
−0.7	Proposed	0.574	0.701	0.889	0.996	1.000	1.000
	D_{NP}	0.098	0.102	0.116	0.264	0.424	0.803
−0.9	Proposed	0.887	0.986	1.000	1.000	1.000	1.000
	D_{NP}	0.124	0.143	0.306	0.698	0.902	0.984

Table 4 Estimates of power and size of 0.05 test of significance for IID noise versus AR(1), $\phi_1 = 0.0, 0.2, 0.4, 0.6$ and 0.8.

$\phi_1^{(2)}$	Method	(n_1, n_2)					
		(100, 50)	(150, 75)	(200, 100)	(500, 250)	(1000, 500)	(2000, 1000)
0.0	Proposed	0.055	0.053	0.053	0.051	0.049	0.050
	D_{NP}	0.054	0.056	0.058	0.052	0.051	0.049
0.2	Proposed	0.428	0.596	0.693	0.908	1.000	1.000
	D_{NP}	0.056	0.058	0.062	0.067	0.077	0.118
0.4	Proposed	0.787	0.924	0.997	1.000	1.000	1.000
	D_{NP}	0.124	0.127	0.163	0.377	0.643	0.884
0.6	Proposed	0.997	1.000	1.000	1.000	1.000	1.000
	D_{NP}	0.321	0.402	0.632	0.918	0.964	0.995
0.8	Proposed	1.000	1.000	1.000	1.000	1.000	1.000
	D_{NP}	0.698	0.834	0.943	0.986	0.996	1.000

where

$$\Sigma = \mathrm{V}_1 + \mathrm{V}_2 \text{ and } \mathrm{V}_i = \begin{bmatrix} (1-(\phi_1^{(i)})^2)(1+\phi_1^{(i)}\theta_1^{(i)}) & -(1-(\phi_1^{(i)})^2)(1-(\theta_1^{(i)})^2) \\ -(1-(\phi_1^{(i)})^2)(1-(\theta_1^{(i)})^2) & (1-(\theta_1^{(i)})^2)(1+\phi_1^{(i)}\theta_1^{(i)}) \end{bmatrix}.$$

We consider $\phi_1^{(1)} = 0.2$ and $\theta_1^{(1)} = -0.5$ for the first population. For the second population, we consider $\phi_1^{(2)} = 0.2$ and $\theta_1^{(2)} = -0.1, -0.3, -0.5, -0.7$ and -0.9.

The two rows opposite of $\theta_1^{(2)} = -0.5$ in Table 3 report the values of $\hat{\alpha}$ for proposed method (above row) and D_{NP} test statistic (bottom row), respectively. Other rows give the values of $\hat{\pi}$ for the other values of $\theta_1^{(2)}$.

Example 4 IID noise versus AR(1).

For the first population, we generate data from IIDN(0, 1). For the second population, we consider the AR(1) model with $\phi_1^{(2)} = 0.0, 0.2, 0.4, 0.6$ and 0.8.

The two rows opposite of $\phi_1^{(2)} = 0.0$ in Table 4 report the values of $\hat{\alpha}$ for proposed method (above row) and D_{NP} test statistic (bottom row), respectively. Other rows give the values of $\hat{\pi}$ for the other values of $\phi_1^{(2)}$.

As can be seen in Tables 1, 2, 3 and 4, the results reveal that the size of two methods is very close to the nominal level ($\alpha = 0.05$), especially as (n_1, n_2) growing, and so the type I error are asymptotically controlled by two methods. On the other hand, the power studies show the excellent discriminatory ability of the proposed method. The proposed test is more powerful than the D_{NP} statistic test. Also, recording the CPU time spent on both methods emphasized the proposed method is computationally intensive.

References

Brockwell PJ, Davis RA (1991) Time series: theory and methods. Springer, Berlin

Caiado J, Crato N, Pena D (2006) A periodogram-based metric for time series classification. Comput Stat Data Anal 50:2668–2684

Camacho M, Perez-Quiroz G, Saiz L (2004) Are European business cycles close enough to be just one? Documento de Trabajo n. 0408, Banco de Espana, Servicio de Estudios

Coates DS, Diggle PJ (1986) Tests for comparing two estimated spectral densities. J Time Ser Anal 7:7–20

Dargahi-Noubary GR (1992) Discrimination between Gaussian time series based on their spectral differences. Commun Stat (Theory and Methods) 21:2439–2458

Diggle PJ, Al Wasel I (1997) Spectral analysis of replicated biomedical time series. Appl Stat 46:31–71

Diggle PJ, Fisher NI (1991) Nonparametric comparison of cumulative periodograms. Appl Stat 40:423–434

Ferguson TS (1996) A course in large sample theory. Chapman & Hall, London

Kakizawa Y, Shumway RH, Taniguchi M (1998) Discrimination and clustering for multivariate time series. J Am Stat Assoc 93:328–340

Maharaj EA (2002) Comparison of non-stationary time series in the frequency domain. Comput Stat Data Anal 40:131–141

Mahmoudi MR, Mahmoodi M (2013a) Inferrence on the ratio of correlations of two independent populations. J Math Ext 7(4):71–82

Mahmoudi MR, Mahmoodi M (2013b) Inferrence on the ratio of variances of two independent populations. J Math Ext 7(2):83–91

Mahmoudi MR, Mahmoudi M, Nahavandi E (2014) Testing the difference between two independent regression models. Commun Stat Theo Meth 45(21):6284–6289