Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v7n1p81

Bayesian analysis of a stationary AR(1) model and outlier

By Kumar et al.

April 26, 2014

Electronic Journal of Applied Statistical Analysis
Vol. 7, Issue 1, 2014, 81-93
DOI: 10.1285/i20705948v7n1p81

Bayesian analysis of a stationary AR(1) model and outlier

Jitendra Kumar*[a], Ashutosh Shukla[b], and Neeraj Tiwari[c]

[a] *Department of Statistics, Central University of Rajasthan, Kishangarh, Rajasthan, India*
[b] *Central Bureau of Health Inteligence, Lucknow, India*
[c] *Departemnt of Statistics, Kumaun University, Almora, Uttrakhand, India*

April 26, 2014

The time varying observation recorded in chronological order is called time series. The extreme values are from the same time series model or appear because of some unobservable causes having serious implications in the estimation and inference. This change deviate the error more and the recorded observation is called outlier. The present paper deals the Bayesian analysis to the extreme value(s) is/are from the same time series model or appears because of some unobservable causes. We derived the posterior odds ratio in different setups of unit root hypothesis. We have also explored the possibility of studying the impact of outlier on stationarity of time series. Using the simulation study, it has been observed that if outlier is ignored a non-stationary series concluded difference stationary.

keywords: Autoregressive model,Outlier, Stationarity, Prior distribution, Posterior odds ratio.

1 Introduction

In data analysis, we deal with large number of sampled or recorded variables. Some of observations are outlying from the rest, which are known as outliers. The time varying observation recorded in chronological order is called time series. The extreme values from the same time series model or appear because of some unobservable causes always have serious implications in the estimation of parameters or testing procedures. Barnett and Lewis (1994) defined that an outlying observation, or outlier, is one that appears

*Corresponding author: jitendra_20932@rediffmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

distinctly from other members of the sample in which it occurs. Main feature of outlier is that it lies outside the overall pattern of a distribution [see,Moore and McCabe (1999)]. In measure studies outliers are often considered as an error or noise, however they may carry important information. Usually, the presence of an outlier indicates some sort of problem. This can be a case that does not fit the model under study or an error in measurement. The presence of outlier is extensively discussed in many fields like network interruption, geographic information systems, clinical trials, voting irregularity analysis, athlete performance analysis, weather predication and other data mining tasks. For details, one may refer to Ruts and Rousseeuw (1996), Fawcett and Provost (1997), Johnson et al. (1998), Penny and Jolliffe (2001),Lu et al. (2003), and Acuna and Rodriguez (2004).

The outlier study is mainly concern to see the impact of outlier on statistical theory as well as in application and ways to identify it. There are several methods to identify the outlier. According to Pyle (1999), a value is considered outlier if it is far away from other values of the same attribute. Barnett and Lewis (1994) first assumed the parameter's distribution and type of expected outliers, which is not realistic for real data mining procedures. The various studies have taken care the impact and way of handling of an outlier. For details, please refer to Knorr and Ng (1997),Williams and Huang (1997), DuMouchel and Schonlau (1998), Knorr et al. (2000), Breunig et al. (2000), Jin et al. (2001), Williams et al. (2002), Hawkins et al. (2002) and Bay and Schwabacher, (2003).

In the field of regression analysis, outlier is not much explored with time series regression. Little efforts have been done by some researchers like McCulloch and Tsay (1994), who had discussed the estimation of parameters under consideration of outlier. Billor and Kiral (2008) carried out an attractive comparative Monte Carlo simulation study to assess the performance of the multiple outlier detection methods and recommended various outlier detection methods in associated conditions. Balke and Fomby (1994) proposed modified sequential test for detection of outlier and achieved better power. He also obtained the asymptotic distribution of the test statistic.Deutsch et al. (1990) , Balke and Fomby (1994) and Tsay (1988) discussed different detection methods in reference of time series. Chen and Liu (1993) estimated the parameters of ARMA model with outlier and saw their impact on estimated parameters. It was shown by Franses and Haldrup (1994) that the presence of outlier affects the limiting distribution of Dicky and Fuller (1979) test. Shin et al. (1996) discussed it for non-seasonal time-series, whereas Haldrup et al. (2005) discussed it for seasonal data, which tend to over reject the unit root hypothesis. Recently, Haldrup et al. (2011) modified the unit root test for the time series contaminated by additive outlier test proposed by Shin *et al.* (1996).Panichkitkosolkul and Niwitpong (2010) also achieved to get improved predictor using multistep-ahead step prediction in the study of AR(1) time series model context of unit root.

The present paper considers the Bayesian analysis of AR(1) time series model contaminated by additive outlier. As the outlier is the observed value(s), is/are deviated more from the fitted model. In AR(1) time series model we have two regression parameters, autoregressive coefficient and intercept term. Obviously the extreme ups and down on the observation will affect the trend, for which we may fit a time series after adjusting the maximum deviated observation (outlier) at which deviation is also maximum in error variable. In such circumstances maximum error is divided into two part (i) observer error from the error distribution and (ii) amplitude of outlier. After partitioning the error, we tested whether the increase in the error is due to the presence of outlier or from the error distribution. The posterior probabilities are obtained in respect to different setups of outlier and unit root hypothesis. Using the posterior probabilities, the posterior odds ratios have been obtained to identify the outlier for a non-stationary and difference stationary series. A simulation study is carried out to take care of the derived theorems. It was observed that the outlier has serious impact on the unit root hypothesis and the posterior distribution has correctly identified the outlier.

2 Model with Intercept Trend and Hypothesis

We consider the time series model

$$y_t = \mu + u_t;\, t\text{=}1,\ 2,\ 3 \ldots T \tag{1}$$

$u_t$is disturbance term and following AR(1) as

$$u_t = \rho u_{t-1} + v_t \tag{2}$$

Where $v_t = \begin{cases} \varepsilon_t & \text{if } t \neq T_{OUT} \\ \lambda e_t + \varepsilon_t & \text{if } t = T_{OUT} \end{cases}$

T_{OUT} is the time point at which series is contaminated by additive outlier. Utilizing the equation (1) and (2) we can write the model

$$y_t = \rho y_{t-1} + (1-\rho)\,\phi + \lambda e_t + \varepsilon_t \tag{3}$$

The time series model (3) is stationary and contaminated by outlier equivalent to the hypothesis $H^1 : \rho \in S, \lambda > 0$ We may rewrite the model (1) under different setups, such as series is stationary and not contaminated by outlier which is equivalent to hypothesis $H^2 : \rho \in S, \lambda = 0$, and the model reduces to

$$y_t = \rho y_{t-1} + (1-\rho)\,\phi + \varepsilon_t \tag{4}$$

Similarly, series is difference stationary and not contaminated by additive outlier hypothesis, which is equivalent to $H^3 : \rho = 1, \lambda = 0$, and the model reduces to

$$\Delta y_t = \varepsilon_t \tag{5}$$

3 Prior Distribution

We take the prior distribution of the parameters of the $\phi \sim N\left(y_0, \frac{1}{(1-\rho^2)\tau}\right)$ $\lambda \sim N\left(\lambda_0, \tau^{-1}\omega\right)$ $p\left(\tau\right) \propto \frac{1}{\tau}$; $0 < \tau < \infty$ $p\left(\rho\right) = \frac{1}{1-a}$; $a < \rho < 1$ The prior odds ratio in favor of H_0 is

$$\frac{p(H_0)}{p(H_1)} = \frac{p_0}{1-p_0} \tag{6}$$

4 Posterior Probabilities

Let us define the following notations for obtaining the posterior probabilities under consideration of hypothesis $H^i; i = 1, 2, 3$

$I = \left[\sum_{t=1}^{T} e_t^2 + \frac{1}{\omega}\right]$

$G(\rho) = \left[T(1-\rho)^2 + \left(1-\rho^2\right)\right]$

$J(\rho) = \left[G(\rho) - \frac{1}{I}\left(\sum_{t=1}^{T} e_t (1-\rho)\right)^2\right]$

$\hat{\phi}_{OUT} = \frac{1}{J(\rho)}\Big[(1-\rho)\sum_{t=1}^{T}(y_t - \rho y_{t-1}) + \left(1-\rho^2\right) y_0$

$-\frac{1}{I}\sum_{t=1}^{T} e_t (1-\rho)\left(\sum_{t=1}^{T} e_t (y_t - \rho y_{t-1}) + \frac{\lambda_0}{\omega}\right)\Big]$

$K(\rho) = \Big[\sum_{t=1}^{T}(y_t - \rho y_{t-1})^2 + \left(1-\rho^2\right) y_0^2 + \frac{\lambda_0^2}{\omega}$

$-\frac{1}{I}\left(\sum_{t=1}^{T} e_t (y_t - \rho y_{t-1}) + \frac{\lambda_0}{\omega}\right)^2 - \hat{\phi}_{OUT}^2 J(\rho)\Big]$

$$H(\rho) = \left[\sum_{t=1}^{T}(y_t - \rho y_{t-1})^2 + \left(1-\rho^2\right) y_0^2 - \hat{\phi}_{OUT}^2 G(\rho)\right] \tag{7}$$

Theorem A

The posterior probability under H^1, in which series is contaminated by outlier is given by

$$P\left(y \mid H^1\right) = \Gamma\left(\frac{T}{2}\right)\frac{1}{2(\pi)^{\frac{T}{2}}\omega^{\frac{1}{2}}(1-a)I^{\frac{1}{2}}}\int_a^1 \frac{\left(1-\rho^2\right)^{\frac{1}{2}}}{J(\rho)^{\frac{1}{2}} K(\rho)^{\frac{T}{2}}} d\rho. \tag{8}$$

Theorem B

The posterior probability under H^2, in which series is not contaminated by outlier is given as

$$P\left(y \mid H^2\right) = \Gamma\left(\frac{T}{2}\right)\frac{1}{(\pi)^{\frac{T}{2}}(1-a)}\int_a^1 \frac{\left(1-\rho^2\right)^{\frac{1}{2}}}{G(\rho)^{\frac{1}{2}} H(\rho)^{\frac{T}{2}}} d\rho \tag{9}$$

Theorem C

The posterior probability under H^3, where we considered that series contains unit root and not contaminated by outlier is

$$P\left(y \mid H^3\right) = \Gamma\left(\frac{T}{2}\right)\frac{1}{(\Delta y_t)^{\frac{T}{2}}} \tag{10}$$

Utilizing the theorems A, B, and C, one can test that series is contaminated by outlier or not, under consideration of series is difference stationary or non stationary.

5 Numerical Illustration

In this section, we obtained the posterior odds ratio in reference of testing the hypothesis regarding different issues of stationarity of time series with contamination of additive outlier. The true model for generated time series of length 40 and intercept term ϕ=100, used different combinations of the parameters δ and ρ. We generated the series by the time series model

$$y_t = \rho y_{t-1} + (1-\rho)\,\phi + \lambda e_t + \varepsilon_t \tag{11}$$

Where $\varepsilon_t \sim N(0,1)$ is the disturbance term ande_t, the amplitude of outlaying observation existed at time T_0. The 500 series generated in different setups of the parameters of time series are mentioned in the Table 1-4. The posterior probabilities given by equation numbers (8), (9) and (10) are calculated for the generated series to obtain the following posterior odds ratios:

1. POR_1 H_0:$\rho = 1, \lambda = 0$, H_1:$\rho \in S, \lambda = 0$, equivalent to series is difference stationary against the alternative that the series is stationary without consideration of outlier, which we tested by the posterior odds ratio given below:

$$\beta_{POR_1} = \frac{p_0}{1-p_0}\frac{\frac{1}{(\Delta y_t)^{\frac{T}{2}}}}{\frac{1}{(1-a)}\int_a^1 \frac{(1-\rho^2)^{\frac{1}{2}}}{G(\rho)^{\frac{1}{2}}H(\rho)^{\frac{T}{2}}}d\rho} \tag{12}$$

2. POR_2 H_0:$\rho = 1, \lambda = 0$, H_1:$\rho \in S, \lambda > 0$ is equivalent to series is difference stationary and not contaminated by outlier against the alternative that the series is stationary with consideration of contamination of outlier, we tested by the posterior odds ratio by

$$\beta_{POR_2} = \frac{p_0}{1-p_0}\frac{\frac{1}{(\Delta y_t)^{\frac{T}{2}}}}{\frac{1}{2\omega^{\frac{1}{2}}(1-a)I^{\frac{1}{2}}}\int_a^1 \frac{(1-\rho^2)^{\frac{1}{2}}}{J(\rho)^{\frac{1}{2}}K(\rho)^{\frac{T}{2}}}d\rho} \tag{13}$$

3. POR_3 H_0:$\lambda = 0; \rho \in S$, H_1:$\lambda > 0, \rho \in S$ is equivalent to the null hypothesis that the series is not contaminated by outlier against the alternative that the series is contaminated by outlier for a stationary series, we can test by

$$\beta_{POR_3} = \frac{p_0}{1-p_0}\frac{\int_a^1 \frac{(1-\rho^2)^{\frac{1}{2}}}{G(\rho)^{\frac{1}{2}}H(\rho)^{\frac{T}{2}}}d\rho}{\frac{1}{2\omega^{\frac{1}{2}}I^{\frac{1}{2}}}\int_a^1 \frac{(1-\rho^2)^{\frac{1}{2}}}{J(\rho)^{\frac{1}{2}}K(\rho)^{\frac{T}{2}}}d\rho} \tag{14}$$

The Tables 1-4 provide posterior odds ratio under different setup of parameters of time series model, the series is generated for λ = -10, -5, 0, 5, 10, ρ=0.90, 0.92, 0.94, 0.96, 0.98 and e_t=25, 50, 75, 100. The null hypothesis H^1 that the series is difference stationary against the alternative that the series is stationary without consideration of outlier is accepted for all setups. In testing the hypothesis H^2 that the series is difference stationary, where outlier is not taken into account against the alternative that the series is stationary with consideration of contamination of outlier is rejected for different setups except for $\lambda=0$, where true series is not contaminated by outlier. In null Hypothesis H^3, under stationary model of time series, it is tested that series is not contaminated by outlier against the alternative that it is contaminated by outlier under consideration of non-stationary series. This may be transferred in a stationary series from the mean of the series and reject the null hypothesis except for the true model. The POR_3 correctly specified the model in all numerical setups.

We found that the unit root hypothesis for all the cases is accepted if outlier is not taken into account and the null hypothesis is rejected at all setups except the true model is not contaminated by outlier. It is also observed that if a series, which is contaminated by additive outlier is not taken into account; the unit root hypothesis may be reversed. For the real data, first identify the time point at which an observed value deviating maximum. Using POR_3 test that this is an outlier or not and then proceed to test the stationarity of the time series. As present study shows that if outlier is not taken into account in the series a non-stationary series may be concluded as difference stationary. In this situation the autoregressive parameter will also be over estimated.

Table 1

λ	**e_t=25**					
	ρ	0.9	0.92	0.94	0.96	0.98
-10	POR_1	2.9829	3.9437	4.9055	4.6343	6.0076
	POR_2	1.08E-52	5.90E-53	1.23E-52	1.23E-51	2.69E-55
	POR_3	6.07E-53	4.72E-54	4.23E-53	2.19E-52	4.80E-56
-5	POR_1	2.9641	3.942	4.8868	4.6038	6.0196
	POR_2	1.31E-36	6.67E-37	1.39E-31	1.21E-27	1.81E-38
	POR_3	7.35E-37	5.58E-38	4.83E-32	2.89E-28	2.38E-39
0	POR_3	24	22.991	12.228	7.1103	20.791
5	POR_1	3.2497	3.888	4.8586	4.7891	6.6476
	POR_2	3.42E-38	6.30E-36	3.63E-38	3.19E-29	6.85E-34
	POR_3	1.38E-38	2.21E-36	1.12E-38	5.48E-30	9.87E-35
10	POR_1	3.2515	3.4855	4.3044	6.1457	6.0871
	POR_2	1.58E-54	9.45E-53	3.22E-55	7.39E-55	7.10E-55
	POR_3	6.97E-55	2.21E-53	7.42E-56	8.30E-56	1.28E-55

Table 2

λ	$\mathbf{e}_t$ **= 50**					
	ρ	0.9	0.92	0.94	0.96	0.98
-10	POR_1	3.2276	3.5134	4.2974	6.0757	6.0668
	POR_2	1.03E-50	9.83E-51	1.78E-52	1.15E-48	1.83E-49
	POR_3	3.50E-51	4.63E-51	5.43E-53	1.70E-49	3.05E-50
-5	POR_1	3.2329	3.5074	4.2998	6.0906	6.0736
	POR_2	4.31E-64	1.72E-62	5.24E-60	1.05E-61	8.22E-60
	POR_3	1.46E-64	6.29E-63	5.80E-61	1.73E-62	1.52E-60
0	POR_3	31.292	5.429	8.6688	22.182	20.374
5	POR_1	3.0145	3.9428	4.9424	4.658	5.9756
	POR_2	1.83E-49	1.21E-51	9.27E-48	4.00E-40	6.44E-52
	POR_3	4.64E-50	7.63E-52	2.04E-48	1.01E-40	1.25E-52
10	POR_1	3.0074	3.9437	4.9327	4.6558	5.9849
	POR_2	8.53E-63	1.05E-63	4.30E-63	2.27E-62	2.02E-63
	POR_3	3.31E-63	3.00E-64	1.44E-63	5.42E-63	5.33E-64

Table 3

Λ	$\mathbf{e}_t$ **=75**					
	ρ	0.9	0.92	0.94	0.96	0.98
-10	POR_1	3.3532	3.871	4.8595	4.8587	6.7184
	POR_2	7.15E-57	2.44E-54	6.67E-55	1.65E-53	1.42E-52
	POR_3	2.32E-57	1.03E-54	1.15E-55	3.24E-54	7.60E-54
-5	POR_1	3.3372	3.8749	4.859	4.8514	6.7155
	POR_2	1.58E-63	2.56E-60	6.55E-63	2.86E-61	4.05E-60
	POR_3	6.70E-64	3.41E-61	1.90E-63	2.20E-62	5.31E-61
0	POR_3	0.28614	22.554	13.166	0.02876	1.9896
5	POR_1	3.2484	3.4906	4.3036	6.1369	6.0855
	POR_2	4.02E-55	5.75E-54	5.70E-57	1.72E-53	1.49E-56
	POR_3	1.48E-55	1.65E-54	1.86E-57	6.26E-55	2.65E-57
10	POR_1	3.2429	3.4959	4.3028	6.1205	6.0816
	POR_2	8.42E-66	7.36E-64	3.73E-60	3.67E-60	2.42E-60
	POR_3	3.83E-66	2.82E-64	4.07E-61	6.13E-61	4.37E-61

Table 4

λ	$\mathbf{e}_t$**=100**					
	ρ	0.9	0.92	0.94	0.96	0.98
-10	POR1	2.9909	3.9443	4.915	4.6399	6.0011
	POR2	4.85E-59	3.14E-59	1.35E-54	1.57E-50	4.36E-61
	POR3	2.73E-59	3.00E-60	4.64E-55	2.78E-51	8.28E-62
-5	POR1	2.8356	3.8105	4.6669	5.2829	6.9437
	POR2	3.52E-66	2.14E-53	3.53E-63	1.47E-65	8.73E-52
	POR3	5.85E-67	7.24E-54	1.06E-63	2.94E-66	1.11E-52
0	POR3	10.41	5.199	27.677	41.432	18.541
5	POR1	3.3047	3.8813	4.8601	4.8307	6.6958
	POR2	1.46E-57	1.10E-57	1.70E-59	8.15E-52	7.88E-56
	POR3	5.55E-58	1.60E-58	1.15E-60	1.49E-52	1.13E-56
10	POR1	3.0036	3.9439	4.9282	4.6532	5.9891
	POR2	3.27E-66	3.45E-66	3.09E-64	5.97E-65	1.88E-65
	POR3	1.29E-66	1.26E-66	1.01E-64	1.64E-65	4.80E-66

Aknowledgment

First author gratefully acknowledge the CST, Uttar Pradesh and SHIATS, Allahabad for Young Scientist visiting fellowship under which the work is carried out.

References

Acuna, E. and Rodriguez, C., (2004). A Meta Analysis Study of Outlier Detection Methods in Classification. *Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez.*

Balke, N.S. and Fomby, T.S.(1994). Large shocks, small shocks, and economic fluctuations: Outliers in macroeconomic time series. *Journal of Applied Econometrics 9, 181–200.*

Bay, S.D. and Schwabacher, M. (2003). Mining distance-based outliers in near linear time with randomization and a simple pruning rule. *Preceding of the ninth ACM-SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA.*

Billor N. and Kiral G. (2008). A Comparison of Multiple Outlier Detection Methods for Regression Data. *Communications in Statistics–Simulation and Computation, 37 (3) 521–545*

Breunig, M. M., Kriegel, H. P., Ng, R. T. and Sander, J. (2000). Identifying density-based local outliers. *In Preceding ACMSIGMOD Conference, 93–104*

Barnett,V. and Lewis,T.(1994). Outliers in Statistical Data. *3rd Edition, John Wiley & Sons*

Chen, C. and Liu, L. (1993). Joint estimation of model parameters and outlier effect in time series. *Journal of American statistical Association, 88, 284-97*

Deutsch, S. J., Richards, J. E. and Swain, J. J.(1990). Effects of a single outlier on ARMA identification. *Communications in Statistics–Theory and Methods, 19 (6) , 2207–2227*

Dickey, D.A., W.A. Fuller(2002). Distribution of the estimators for autoregressive time series with unit root. *journal of the American Statistical Association,74,427-431 (at page 108 in third paragraph line 13)*

DuMouchel, W. and Schonlau, M. (1998). A fast computer intrusion detection algorithm based on hypothesis testing of command transition probabilities. *In Proceedings of the 4th International Conference on Knowledge Discovery and Data-mining (KDD98), 189–193*

Fawcett, T. and Provost, F. (1997). Adaptive fraud detection. *Data-mining and Knowledge Discovery, 1, 291–316*

Franses, P. and Haldrup, N. (1994). The effects of additive outliers on tests for unit roots and cointegration. *Journal of Business and Economic Statistics 12, 471-478*

Haldrup, N., Montas, A. and Sans, A. (2005). Measurement errors and outliers in seasonal unit root testing. *Journal of Econometrics, 127, 103-128.*

Haldrup, N., Montanes, A. and Sanso, A. (2011). Detection of outlier in seasonal time series. *Journal of Time Series Econometrics, 3, article 2.*

Hawkins, S., Williams, H. X. and Baxter, R. A. (2002). Outlier detection using replicator neural networks. *In Proceedings of the Fifth International Conference and Data Warehousing and Knowledge Discovery (DaWaK02), Aix en Provence, France.*

Jin, W., Tung, A. and Han, J. (2001). Mining top local outliers in large databases. *In Proceedings of the 7th International Conference on Knowledge Discovery and Data-mining (KDD), San Francisco, CA.*

Johnson, T., Kwok, I. and Ng, R. (1998). Fast Computation of 2-Dimensional Depth Contours. *In Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD),AAAI Press, 224–228*

Knorr, E. and Ng, R. (1997). A unified approach for mining outliers. *In Proceedings Knowledge Discovery and Data-mining (KDD), 219-222*

Knorr, E., Ng, R. and Tucakov, V. (2000). Distance-based outliers: Algorithms and applications. *VLDB Journal: Very Large Data Bases, 8(3-4), 237-253.*

Lu, C., Chen, D. and Kou, Y. (2003). Algorithms for spatial Outlier detection. *In Proceedings of the 3rd IEEE International Conference on Data-mining (ICDM'03), Melbourne, FL.*

McCulloch, R. E. and Tsay, R. S. (1994). Bayesian analysis of autoregressive time series via the Gibbs Sampler. *Journal Time Series Analysis, 15, 235-50.*

Moore, D. S. and McCabe, G. P. (1999). Introduction to the Practice of Statistics. *third edition New York: W. H. Freeman.*

Panichkitkosolkul, W. and Niwitpong, S. A. (2010). Multistep-ahead predictors for Gaussian AR(1)process with additive outlier following unit root tests. *International Journal of Innovative Management, Information & Production, 1, 49-55*

Penny, K. I. and Jolliffe, I. T. (2001). A comparison of multivariate Outlier detection methods for clinical laboratory safety data. *The Statistician, 50, 295-308*

Pyle, D. (1999). Data Preparation for Data Mining. *Morgan Kaufmann Publishers, Los Altos, California*

Ruts, I. and Rousseeuw, P. (1996). Computing Depth Contours of Bivariate Point Clouds. *In Computational Statistics and Data Analysis, 23,153-168*

Shin, D. W., Sarkar, S. and Lee, J. H.(1996). Unit root tests for time series with outliers. *Statistics and Probability Letters, 30, 189-197*

Tsay, R. S. (1988). Outlier, level shift, and variance changes in time series. *Journal of Forecasting, 7, 1-20*

Williams, G. J. and Huang, Z. (1997). Mining the knowledge mine: The hot spots methodology for mining large real world databases. *In Abdul Sattar, editor, Advanced Topics in Artificial Intelligence, volume 1342 of Lecture Notes in Artificial Intelligence, 340-348, Springer*

Williams, G. J., Baxter, R. A., He, H. X., Hawkins, S. and Gu, L. (2002). A Comparative Study of RNN for Outlier detection Data Mining. *In EEE International Conference on Data-mining (ICDM'02), Maebashi City, Japan, CSIRO Technical Report CMIS-02/102.*

Appendix
Derivation of Theorem A

The likelihood function under the condition that series is stationary and contaminated by outlier:

$$\begin{aligned} p\left(y|H^1\right) &= L\left(y|\phi,\tau,\rho,\lambda\right)p\left(\phi|\tau\rho\right)p\left(\lambda|\tau\rho\right)p\left(\tau\right)p\left(\rho\right) \\ &= \int_a^1\int_0^\infty\int_{-\infty}^\infty\int_0^\infty \frac{\tau^{\frac{T}{2}}\left(1-\rho^2\right)^{\frac{1}{2}}}{\left(2\pi\right)^{\frac{T}{2}+1}\omega^{\frac{1}{2}}\left(1-a\right)} exp\left[-\frac{\tau}{2}\left\{\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-\left(1-\rho\right)\phi-\lambda e_t\right)^2\right.\right. \\ &\left.\left. +\frac{\left(\lambda-\lambda_0\right)^2}{\omega}+\left(1-\rho^2\right)\left(\phi-y_0\right)^2\right\}\right]d\lambda d\phi d\tau d\rho \qquad (A_1) \\ &= \int_a^1\int_0^\infty\int_{-\infty}^\infty\int_0^\infty \frac{\tau^{\frac{T}{2}}\left(1-\rho^2\right)^{\frac{1}{2}}}{\left(2\pi\right)^{\frac{T}{2}+1}\omega^{\frac{1}{2}}\left(1-a\right)} exp\left[-\frac{\tau}{2}\left\{\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-\left(1-\rho\right)\phi\right)^2+\lambda^2\sum_{t=1}^{T}\left(e_t\right)^2\right.\right. \\ &\left.\left. -2\lambda\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-\left(1-\rho\right)\phi\right)^2 e_t+\frac{\lambda^2}{\omega}+\frac{{\lambda_0}^2}{\omega}-\frac{2\lambda\lambda_0}{\omega}+\left(1-\rho^2\right)\left(\phi-y_0\right)^2\right\}\right]d\lambda d\phi d\tau d\rho \end{aligned}$$

Now taking

$$\lambda^2\sum_{t=1}^{T}\left(e_t\right)^2-2\lambda\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-\left(1-\rho\right)\phi\right)^2 e_t+\frac{\lambda^2}{\omega}+\frac{{\lambda_0}^2}{\omega}-\frac{2\lambda\lambda_0}{\omega}+\left(1-\rho^2\right)\left(\phi-y_0\right)^2=\left(\lambda-\hat{\lambda}\right)^2 I-\hat{\lambda}^2 I$$

Where

$$\hat{\lambda}=\frac{1}{I}\left[\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-\left(1-\rho\right)\phi\right)\right] \qquad (A_2)$$

Then

$$
\begin{aligned}
p\left(y|{H_1}^1\right) &= \int_a^1\int_0^\infty\int_{-\infty}^\infty\int_0^\infty \frac{\tau^{\frac{T}{2}}\left(1-\rho^2\right)^{\frac{1}{2}}}{(2\pi)^{\frac{T}{2}+1}\,\omega^{\frac{1}{2}}\,(1-a)} exp\left[-\frac{\tau}{2}\left\{\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-(1-\rho)\,\phi\right)^2+\frac{{\lambda_0}^2}{\omega}\right.\right. \\
&\left.\left. +\left(\lambda-\hat{\lambda}\right)^2 I-\hat{\lambda}^2 I+\left(1-\rho^2\right)(\phi-y_0)^2\right\}\right] d\lambda d\phi d\tau d\rho \qquad (A_3) \\
&= \int_a^1\int_0^\infty\int_{-\infty}^\infty \frac{\tau^{\frac{T}{2}}\left(1-\rho^2\right)^{\frac{1}{2}}}{(2\pi)^{\frac{T+1}{2}}\,\omega^{\frac{1}{2}}\,(1-a)\,I^{\frac{1}{2}}}\left(\frac{1}{2}\right) exp\left[-\frac{\tau}{2}\left\{\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}-(1-\rho)\,\phi\right)^2+\frac{{\lambda_0}^2}{\omega}\right.\right. \\
&\left.\left. +\left(1-\rho^2\right)(\phi-y_0)^2-\hat{\lambda}^2 I\right\}\right] d\phi d\tau d\rho \\
&= \int_a^1\int_0^\infty\int_{-\infty}^\infty \frac{\tau^{\frac{T}{2}}\left(1-\rho^2\right)^{\frac{1}{2}}}{(2\pi)^{\frac{T+1}{2}}\,\omega^{\frac{1}{2}}\,(1-a)\,I^{\frac{1}{2}}} exp\left[-\frac{\tau}{2}\left\{\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}\right)^2-T\,(1-\rho)^2\,\phi^2\right.\right. \\
&-2\,(1-\rho)\,\phi\sum_{t=1}^{T}\left(y_t-\rho y_{t-1}\right)+\frac{{\lambda_0}^2}{\omega}+\left(1-\rho^2\right)\phi^2+\left(1-\rho^2\right){y_0}^2-2\left(1-\rho^2\right)\phi y_0 \\
&-\frac{1}{I}\left(\sum_{t=1}^{T}e_t\left(y_t-\rho y_{t-1}\right)+\frac{\lambda_0}{\omega}\right)^2-\frac{1}{I}\left(\sum_{t=1}^{T}e_t\,(1-\rho)\,\phi\right)^2 \\
&\left.\left. +\frac{2}{I}\sum_{t=1}^{T}e_t\,(1-\rho)\,\phi\left(\sum_{t=1}^{T}e_t\left(y_t-\rho y_{t-1}\right)+\frac{\lambda_0}{\omega}\right)\right\}\right] d\phi d\tau d\rho
\end{aligned}
$$

Now we take

$$T(1-\rho)^2\phi^2 - 2(1-\rho)\phi\sum_{t=1}^{T}(y_t-\rho y_{t-1}) + \left(1-\rho^2\right)\phi^2 - 2\left(1-\rho^2\right)\phi y_0 - \frac{1}{I}\left(\sum_{t=1}^{T}(1-\rho)\phi\right)^2 +$$

$$+\frac{2}{I}\sum_{t=1}^{T}e_t(1-\rho)\phi\left(\sum_{t=1}^{T}e_t(y_t-\rho y_{t-1})+\frac{\lambda_0}{\omega}\right) = \left(\phi-\hat{\phi_{OUT}}\right)^2 J(\rho) - \hat{\phi_{OUT}^2}J(\rho) \qquad (A_4)$$

then we get

$$\begin{aligned} p\left(y|{H_1}^1\right) &= \int_a^1\int_0^\infty\int_{-\infty}^{\infty}\frac{\tau^{\frac{T}{2}}\left(1-\rho^2\right)^{\frac{1}{2}}}{(2\pi)^{\frac{T+1}{2}}\omega^{\frac{1}{2}}(1-a)I^{\frac{1}{2}}}exp\left[-\frac{\tau}{2}\left\{\sum_{t=1}^{T}(y_t-\rho y_{t-1})^2\right.\right. \\ &\quad +\frac{{\lambda_0}^2}{\omega}+\left(1-\rho^2\right)\phi^2+\left(1-\rho^2\right){y_0}^2-2\left(1-\rho^2\right)\phi y_0 \\ &\quad -\frac{1}{I}\left(\sum_{t=1}^{T}e_t(y_t-\rho y_{t-1})+\frac{\lambda_0}{\omega}\right)^2 \\ &\quad \left.\left.+\left(\phi-\hat{\phi_{OUT}}\right)^2 J(\rho)-\hat{\phi_{OUT}^2}J(\rho)\right\}\right]d\phi d\tau d\rho \\ &= \int_a^1\frac{\left(1-\rho^2\right)^{\frac{1}{2}}\Gamma\left(\frac{T}{2}\right)}{2(2\pi)^{\frac{T}{2}}\omega^{\frac{1}{2}}(1-a)I^{\frac{1}{2}}J(\rho)}\left(\frac{2}{K(\rho)}\right)^{\frac{T}{2}}d\rho \end{aligned} \qquad (A_5)$$

Similar the posterior probability under $H^2 and H^3$

$$p\left(y|H^2\right) = \Gamma\left(\tfrac{T}{2}\right)\frac{1}{(\pi)^{\frac{T}{2}}(1-a)}\int_a^1\frac{\left(1-\rho^2\right)^{\frac{1}{2}}}{G(\rho)^{\frac{1}{2}}H(\rho)^{\frac{1}{2}}}d\rho \qquad (A_6)$$

$$p\left(y|H^3\right)\Gamma\left(\tfrac{T}{2}\right)\frac{1}{(\pi)^{\frac{T}{2}}(\Delta y_t)^{\frac{T}{2}}} \qquad (A_7)$$

Using equation $(A_5), (A_6) and (A_7)$, we get the desired posterior probabilities in theorem 1A, 1B, and 1C.