# MACROS FOR WORKING WITH NONSTATIONARY TIME SERIES

Ashok Mehta, Ph.D.
AT&T

## Introduction

In time series analysis it is assumed that the underlying stochastic process that generated the series is invariant with respect to time. If the characteristics of the stochastic process are changing over time then the process is called nonstationary. It is extremely difficult, if not impossible, to model a nonstationary time series over past and future intervals of time. However, if a time series is stationary then one can model the process and forecast the series.

Most economic time series (GNP, price index, interest rate, exchange rate, etc.) are nonstationary. Fortunately, many of them can be rendered stationary by simple transformations. This paper demonstrates the use of SAS® (version 6.08 for Windows) for identifying a nonstationary time series and transforming it into a stationary time series. Section I formally defines the stationary time series process. Section II outlines a method for detecting and eliminating nonstationarity using SAS. In Section III an example is given. Finally, in Section IV some drawbacks of the method are discussed. The SAS code and output from the example are given in the Appendix.

## I. Stationary Process

Let $\{y_1, y_2, \ldots, y_T\}$ be a stochastic time series. The series $y_t$ is stationary if the mean, variance and lag covariance of the series are constant over time. In other words, following conditions are fulfilled:

1. Constant Mean:
$$E(y_t) = \mu_y$$

2. Constant Variance:
$$E(y_t - \mu_y)^2 = \sigma_y^2$$

3. Constant Lagged Covariance:
$$E[(y_t - \mu_y)(y_{t+k})] = \nu_k$$

## II. Detection of Nonstationarity

Nonstationarity in a time series can be detected by following these steps:

1. Visualize the series by plotting it. If there is any visible trend or unusual variability then that will suggest existence of nonstationarity.

2. Examine the autocorrelation function (ACF) of the series. A gradual decline in the ACF is also indication of nonstationarity.

3. Apply the Dickey-Fuller test. The Dickey-Fuller test involves a regression of $\Delta y_t$ on $y_{t-1} - \bar{y}$, $\Delta y_{t-1}, \ldots\ldots,$ $\Delta y_{t-p}$, where $\Delta$ indicates a first difference. The number of $\Delta y_{t-p}$ terms included in the regression is one less than the autoregressive process (AR) used to model the series. The $t$ statistics for the parameter estimate associated with $(y_{t-1} - \bar{y})$ provides a statistical test for the stationarity hypothesis (also referred as unit root test). The distribution of $t$ statistic is not like the Student's $t$ distribution but as given by Fuller (1976).

In SAS/ETS® (Version 6.08 for Windows) there are two macros called DFTEST and LOGTEST. These macros are very useful in testing for nonstationarity and eliminating it by simple transformations on the series. The DFTEST macro performs the Dickey-Fuller test. The DFTEST macro computes the $t$ statistic and compares it with the distribution given by Fuller (1976) and provides the correct p-value. The LOGTEST macro examines the series to determine if the logarithmic transformation along with differencing of the series will help in modelling the series. It computes the log likelihood value of the logarithmic transformed and differenced series and compares it with the original differenced series.

In short, the use of the above mentioned macros will identify the existence of nonstationarity and will recommend if differencing and log transformation will make the series stationary.

The DFTEST and LOGTEST macros are modified for this paper to make their output more readily interpretable. The SAS code used for modification is provided in the Appendix.

## III. Example

In order to illustrate the use of concepts discussed above, the Canadian exchange rate (Canadian $/ US $) time series (Jan. 1980 to Mar. 1994) is examined for the existence of nonstationarity. The SAS code and output are given in the Appendix.

Following is the summary of the output given in the Appendix.

1. In Figure 1 the plot of the original series is shown. The plot clearly indicates that the series has a time trend and there is considerable variability in the series. Therefore, the series is likely to have a nonstationarity problem.

2. The estimated autocorrelation function (ACF) for the series shows a gradual decay. This strengthens the possibility of nonstationarity in the series.

3. The modified DFTEST macro is applied to the series. The Dickey-Fuller test finds the series nonstationary. The test is repeated with the first differenced (Note: option DIF=(1,0,0) in the SAS code) series. This time the Dickey-Fuller test rejects the hypothesis of nonstationarity. Therefore, the Canadian exchange rate series is nonstationary but the first differenced series is stationary and suitable for the modelling purposes.

4. Next, the modified LOGTEST macro is applied to the series. The output recommends that the log transformation along with the first difference of the series maximizes the log likelihood value compared to the non-transformed differenced series. Therefore, the log transformation and differencing of the series makes the series stationary and most appropriate for modelling.

5. Figure 2 shows the plot of the log transformed and differenced series. Note there is no time trend and variability is also more stable.

6. Finally, the ACF of the log transformed and differenced series is estimated. There is a sudden decline in the ACF plot compared to the ACF plot of the original series. This confirms stationarity in the transformed series.

## IV. A Word of Caution

The above method suggests that log transformation and differencing of a series can eliminate nonstationarity. There are two drawbacks in implementing the suggestion: (1) By differencing a series one loses observations, which may be a problem with a smaller data series. (2) If a series is mistakenly differenced by a higher order than required then it leads to the problem of over-differencing. Over-differencing will lead to a model that does not represent accurately the true underlying structure of the time series.

## Trademarks



## References

Dickey, D. A. and Fuller, W. A. (1979) "Distribution of the Estimators for Autoregressive Time Series with Unit Root," *Journal of the American Statistical Association*, 74, p. 427-431.

Fuller, W. A. (1976) *Introduction to Statistical Time Series*, John Wiley.

Harvey, A (1989) *The Econometric Analysis of Time Series*, The MIT Press.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lutkepohl, H. and Lee T. (1988) *Introduction to the Theory and Practice of Econometrics*, John Wiley and Sons.

Pindyck, R. S. and Rubinfeld, D. L. (1991) *Econometric Models & Economic Forecasts*, McGraw Hill.

SAS/ETS® User's Guide, Version 6, SAS Institute Inc.

Forecasting Techniques Using SAS/ETS® Software Course Notes, SAS Institute Inc.

SAS/ETS® Software: Application Guide 1, (1992), SAS

Institute Inc.

SAS Guide to Macro Processing, Version 6, SAS Institute Inc.

---

**Acknowledgements**

The Author would like to thank Mr. Bruce McNevin and Mr. Nikos Tsangaroulis of AT&T for their support and encouragement.

Ashok Mehta, Ph.D.
AT&T
295 N. Maple Ave.
Basking Ridge, NJ 07920

Tel: (908) 221-5817
Email: amehta@attmail.com

## Appendix

Below is the code for modifying the DFTEST macro. Insert this code at the end of DFTEST macro (just before %mend).

```
/******************************************************************/
DATA _NULL_;
  FILE PRINT;
  PVALUE=SYMGET('dftest');
  IF PVALUE >0.05 THEN DO;
  PUT //"P-value for the Dickey-Fuller test for variable &var is" @56
  PVALUE 6.4
       @62 '.'
    // 'p-value is greater than 0.05 suggesting nonstationarity'
    /'and the need to difference the series.';
  END;

  ELSE DO;
  PUT //"P-value for the Dickey-Fuller test for variable &var is" @56
  PVALUE 6.4
       @62 '.'
    // 'p-value is lower than 0.05 suggesting stationarity'
    / 'and there is no need to difference the series.';
  END;
RUN;

/******************************************************************/
```

Below is the code for modifying the LOGTEST macro. Insert this code at the end of LOGTEST macro (just before %mend).

```
/******************************************************************/
DATA _NULL_;
  FILE PRINT;
  TRANS=SYMGET('LOGTEST');
  PUT / 'The LOGTEST recommends transformation: '   TRANS;
TITLE "LOGTEST FOR &VAR";

RUN;

/******************************************************************/
```

Following is the code for examining nonstationarity in a time series.

```
/******************************************************************/
OPTIONS LS=65 NODATE PAGENO=1;

FILENAME IN 'A:\EXCAUS';
```

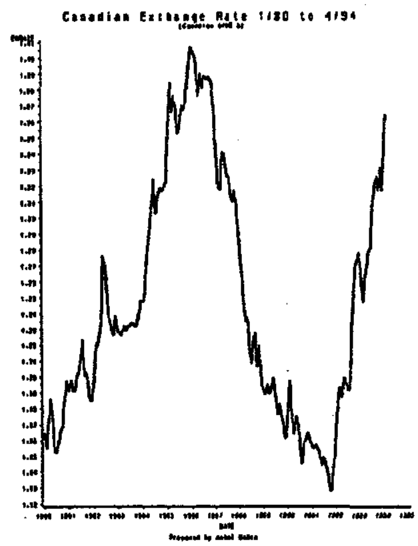
```
DATA EXCAUS;
   INFILE IN FIRSTOBS=8;
   INPUT YYMM EXRATE;
   LEXRATE = LOG(EXRATE);
   LEXRATED = DIF(LEXRATE);
   DATE = INTNX('MONTH', '01JAN1980'D, _N_-1);
   DROP YYMM;

/************************* FIGURE 1 *****************************/
FILENAME OUT1 'A:\CHART1.CGM';

GOPTIONS DEV=CGM ROTATE=LANDSCAPE GSFNAME=OUT1 GSFMODE=REPLACE;
SYMBOL1 I=J W=3;

PROC GPLOT DATA=EXCAUS;
    PLOT EXRATE*DATE /HAXIS = '01JAN1980'D TO '01JAN1995'D BY
YEAR;
    FORMAT DATE YEAR.;
TITLE J=C 'Canadian Exchange Rate 1/80 to 4/94';
TITLE2 J=C '(Canadian $/US $)';
FOOTNOTE J=C 'Prepared by Ashok Mehta';


/****************************** ACF****************************/

PROC ARIMA DATA=EXCAUS;
   I VAR=EXRATE;
TITLE 'IDENTIFYING THE SERIES';

/************************* DICKEY-FULLER TEST*****************/
TITLE 'DICKEY FULLER TEST: ORIGINAL SERIES';
%DFTEST(EXCAUS, EXRATE, DLAG=1,OUT=OUT,TREND=2,OUTSTAT=DSTAT);


TITLE 'DICKEY FULLER TEST: DIFFERENCED SERIES';
%DFTEST(EXCAUS, EXRATE, DLAG=1, DIF=(1,0,0), OUT=OUT, TREND=2,
        OUTSTAT=DSTAT);

/*********************** LOGTEST ********************************/
TITLE 'LOGTEST: ORIGINAL SERIES';

%LOGTEST(EXCAUS, EXRATE,OUT=OUT);

/****************************FIGURE 2****************************/
FILENAME OUT2 'A:\CHART2.CGM';

GOPTIONS DEV=CGM ROTATE=LANDSCAPE GSFNAME=OUT2 GSFMODE=REPLACE;
SYMBOL1 I=J W=3;

PROC GPLOT DATA=EXCAUS;
    PLOT LEXRATED*DATE /HAXIS = '01JAN1980'D TO '01JAN1995'D BY
YEAR ;
```

```
FORMAT DATE YEAR2.;
TITLE J=C '1st DIFF. of LOG(EX. RATE) ';
FOOTNOTE J=C 'Prepared by Ashok Mehta';

/***********************ACF FOR TRANSFORMED SERIES *************/

PROC ARIMA DATA=EXCAUS;
   I VAR=LEXRATED;
TITLE 'STATIONARY SERIES';
RUN;
/***************************************************************/
```

Following is the SAS output:

**OUTPUT**

**Figure 1:**




```
DICKEY FULLER TEST: ORIGINAL SERIES                1

             ARIMA Procedure

  Name of variable = EXRATE.

  Mean of working series = 1.247304
  Standard deviation     = 0.076437
  Number of observations =      171
```

```
                          Autocorrelations

Lag Covar   Corr    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0 0.00584 1.00000 |                    |********************|
  1  0.0057 0.97588 |                 .  |********************|
  2 0.00554 0.94807 |               .    |******************* |
  3  0.0054 0.92463 |             .      |******************  |
  4 0.00525 0.89936 |            .       |******************  |
  5 0.00511 0.87421 |           .        |*****************   |
  6 0.00496 0.84849 |           .        |*****************   |
  7  0.0048 0.82149 |          .         |****************    |
  8 0.00464 0.79451 |         .          |****************    |
  9 0.00447 0.76567 |         .          |***************     |
 10 0.00428 0.73267 |        .           |***************     |
 11 0.00406 0.69541 |        .           |**************      |
 12 0.00383 0.65536 |       .            |*************       |
 13  0.0036 0.61643 |       .            |************.       |
 14 0.00338 0.57922 |       .            |************.       |
 15 0.00316 0.54027 |       .            |*********** .       |
 16 0.00291 0.49863 |      .             |**********   .      |
 17 0.00267 0.45672 |      .             |*********    .      |
 18 0.00242 0.41390 |      .             |********     .      |
 19 0.00216 0.36932 |      .             |*******      .      |
 20 0.00193 0.32977 |      .             |*******      .      |
 21 0.00169 0.28962 |      .             |******       .      |
 22 0.00144 0.24691 |      .             |*****        .      |
 23 0.00118 0.20167 |      .             |****         .      |
 24 0.00093 0.15933 |      .             |***          .      |
                             "." marks two standard errors
```

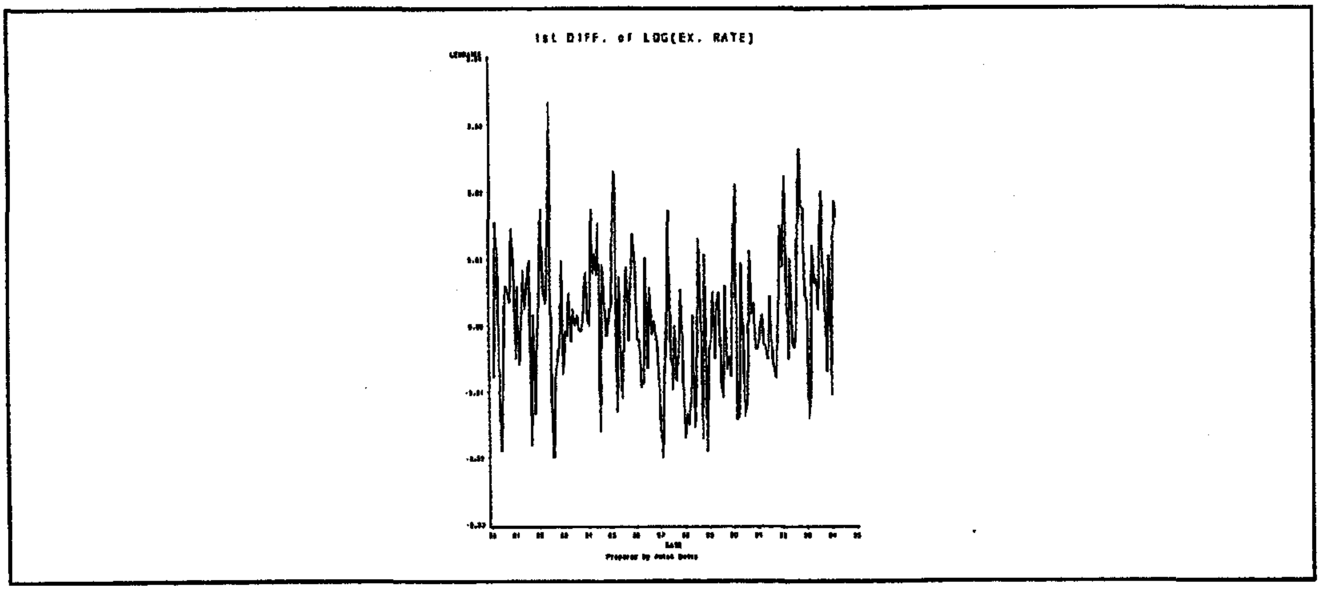
---

Output has been truncated to save space.

---

```
                DICKEY FULLER TEST: ORIGINAL SERIES                 3

P-value for the Dickey-Fuller test for variable EXRATE 0.7322.
p-value is greater than 0.05 suggesting nonstationarity
and the need to difference the series.

              DICKEY FULLER TEST: DIFFERENCED SERIES                4

P-value for the Dickey-Fuller test for variable EXRATE 0.01   .
p-value is lower than 0.05 suggesting stationarity
and there is no need to difference the series.

                        LOGTEST FOR EXRATE                          5

The LOGTEST recommends transformation: LOG
```

**Figure 2**




```
                         STATIONARY SERIES                         6

                          ARIMA Procedure

                Name of variable = LEXRATED.

                Mean of working series = 0.000933
                Standard deviation     = 0.010111
                Number of observations =      170

                         Autocorrelations

Lag Covar    Corr     -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  0  0.0001 1.00000   |                    |********************|
  1 0.00002 0.19178   |                  . |****                |
  2   -15E-7 -.01472  |                  . |   .                |
  3 3.82E-6 0.03739   |                  . |*  .                |
  4 2.32E-6 0.02273   |                  . |   .                |
  5 -988E-8 -.09664   |                  .**|   .               |
  6 4.92E-6 0.04814   |                  . |*  .                |
  7 7.04E-6 0.06889   |                  . |*  .                |
  8 0.00002 0.19662   |                  . |****                |
  9 0.00002 0.16766   |                  . |***                 |
 10 0.00002 0.23261   |                  . |*****               |
 11 0.00002 0.15318   |                 .  |***.                |
                 .    |   .                |
                         "." marks two standard errors
```

Output has been truncated to save space.