Journal of the Nigerian Association of Mathematical Physics
Volume **15 (November, 2009), pp 101 - 106**

Some basic tests on time series outliers

N. P. Olewuezi
Department of Statistics,
Federal University of Technology Owerri, Nigeria.

Abstract

Outliers are common place in applied time series analysis and various types of structural changes occur frequently and raises the question of efficiency and adequacy in fitting models. The methods under consideration for the tests of time series outliers are the Peirce's criterion, Chauvenet's criterion and Grubbs' test. A set of data was considered and later on tested for outliers. From the findings, the Peirce's criterion identified two outliers in the data set while the Chauvenet's and Grubbs' tests both identified only one outlier. In the Peirce's criterion, the result of two outliers were opposed by the Chauvenet's criterion and Grubb's Test because Peirce's criterion accounts for the case where there is more than one suspect data point at once.

Keywords

Time series analysis; outliers; data point, criterion.

1.0 Introduction

Outlier detection method has no exact definition. An exact definition often depends on hidden assumptions regarding the data structure and the detection method. There are some general definitions that cope with various types of data and methods. Observations that deviate from the rest of the observations exist frequently in data.

Hawkins [7] defined outliers as observations that deviate so much from other observations in the same data set as to arouse suspicions that it was generated by a different mechanism. Barnet and Lewis [2] defines an outlying observation as one that appears to deviate markedly from other members of the sample in which it occurs. Other case-specific definitions are given below.

Possible sources of outlier are recording and measurement errors, incorrect distribution assumptions and unknown data structure. Studies have shown that data sets containing outliers are always misleading. Rasmussen [10] showed that outliers can increase error variance and reduce the power of statistical tests.

Harvey, [6] in his work argued that although outliers make statistics difficult, before such an outlier is deleted or retained, the analyst must ask himself the following questions:

(i) was the value entered into the data set correctly?
(ii) were there any experimental problems with such a value?
(iii) is the outlier caused by biological diversity?

If the answers are no, Harvey [6] suggested that:

(a) the outlier was due to chance. Here, the researcher should retain the value in the analysis.
(b) the outlier was due to a mistake and thus should be deleted from the analysis.

e-mail address: ngolewe@yahoo.com
Telephone: 234-08034933133

Osborne and Overbay [9] in their work brought to light the power of any outlier(s) in data analysis (no matter the cause) and advised on instant detection because the importance of accuracy in any analysis cannot be over emphasized.

1.1 Types of time series outliers

In time series analysis, outliers can take several forms. Fox [4] proposed a classification of time series outliers to type I and type II based on the autoregressive model. These two types were later renamed as additive and innovational outliers (AO and IO). Tsay [13] defined three other types of outliers as Level Shift (LS), (sometimes called Level Change, LC), Transient Change (TC) and Variance Change (VC).

1.1.1 Additive outlier (AO)

An additive outlier affects a single observation which is either larger or smaller in value than expected. After this disturbance, the series returns to its normal path as if nothing had happened.

The AO model is given by

$$Z_t = \begin{cases} X_t\,, & t \neq \tau \\ X_t + \omega_A, & t = \tau \end{cases} \tag{1.1}$$

where X_t = Outlier Free Series, Z_t = Observed Series, τ = Time at which the outlier occurs, ω_A = Magnitude of the outlier. Equation (1.1) can be written as:

$$Z_t = X_t + \omega_A I_t^{(\tau)} = \frac{\theta(B)}{\phi(B)} a_t + \omega_A I_t^{(\tau)}$$

where $I_t^{(\tau)} = \begin{cases} 1, & t = \tau \\ 0, & t \neq \tau \end{cases}$ is an indicator variable which is zero at all except at time $t = \tau$. Equivalently,

$$a_t = \Pi(B)\, Z_t - \omega_A I_t^{(\tau)}\ .$$

1.1.2 Innovational outliers (IO)

This affects several observations and it corresponds to an internal error. An IO model is given by: $Z_t = X_t + \dfrac{\theta(B)}{\phi(B)}\omega_L I_t^{(\tau)} = \dfrac{\theta(B)}{\phi(B)}\left(a_t + \omega_L I_t^{(\tau)}\right)$, equivalently, $a_t = \Pi(B) Z_t - \omega_L I_t^{(\tau)}$

In any ARMA model, the estimated residuals is given by

$$e_t = a_t + \omega_L\,,\ t = \tau$$
$$e_t = a_t,\ t = \tau + j$$

For large samples, its effects can be neglected.

1.1.3 Level shift (LS) or Level change (LC)

It simply changes the level (or mean) of the series by a certain magnitude ω, from a certain observation onwards. It can be seen as sequence of additive outlier of the same size.

$$Z_t = \begin{cases} X_t, & t < \tau \\ X_t + \omega_L, & t \geq \tau \end{cases}$$

The model changes from Z_t to X_t to $Z_t = X_t + \omega_L$, where ω_L can be either positive or negative.

1.1.4 Transient change (TC) or Temporary change (TC)

This is a generalization of additive outlier and level shift in the sense that it causes an initial impact like an additive outlier but the effect is passed on the following observations. Meanwhile, the impact of a TC is not permanent, however, it decays exponentially. Formally, a TC has an initial effect of $\omega_o = \omega_T$ and this effect dies out gradually with $\dfrac{\omega(B)}{\sigma(B)} = \dfrac{1}{(1-\sigma B)}$, where σ is the rate of decay, $0 < \sigma < 1$.

Note that with the limit σ = 0, this becomes an additive outlier with σ = 1, a level shift. This will in the same model be very close to an IO.

1.1.5 Variance change (VC)

A VC does not affect the level of the series directly like the other types considered. It is still further away from the additive and innovational outlier types and is not usually considered in connection with outlier at all. It simply changes the variance of the observed series by a new zero mean.

2.0 Some methods of identification of outliers

The methods to be examined here include

(i) Peirce's criterion, (ii) Chauvenet's criterion and (iii) Grubbs' test

2.1 Peirce's criterion

Benjamin, [3], addressed the problem of identifying outliers. He used the elements of probability theory to develop a logical method to eliminate suspicious data from a data set. Ross, [11], stated in his work that Peirce's Criterion was formulated based on the principle that the proposed observations should be rejected when the probability of system of outliers obtained by retaining them is less than that of the system of outliers obtained by their rejection, multiplied by the probability of making so many, and no more abnormal observations. The actual method of calculation used by Peirce was mathematically cumbersome. Ross, [11], gave the basic steps used by Peirce in the derivation of this criterion as:

(i) obtain the equation for the probability of occurrence of a particular deviation from the mean based on the normal distribution. Consider the mean to be based on all the data values with none rejected.

(ii) use the result from step (i) to obtain the probability of the system for all deviations with the condition that a that a deviation limit (to be determined) is exceeded.

(iii) Obtain the probability of the system of deviation of the data set with some data points removed.

(iv) Allow rejection of the data values if the probability from step (i) is less than step (ii)

(v) From the inequality in step (iv) determine the ratio of the limiting deviation to the standard deviation of the sample data. The resulting ratio is what is recorded in Peirce's table as a function of the number of observations or data values and the number of suspect data values.

The Peirce's method of identifying outliers uses the following procedures:

(a) calculate the mean and sample standard deviation of the complete data set.

(b) Obtain an R-value from Peirce's table corresponding to the number of collected measurements. Assume the case of one outlier first, even if there appears to be more than one.

(c) Calculate $|x_i – \text{mean}|_{\max}$ = R-value x (standard deviation) and $|x_i – \text{mean}|$

If $|x_i – \text{mean}| > |x_i – \text{mean}|_{\max}$, then the data point x_I is an outlier.

(d) If only one outlier is identified as a result of step (b), assume the case of two outliers while maintaining the original values of the mean, standard deviation and original number of measurements. Then go to step (e)

(e) If more than one outlier is found as a result of step (b), assume the next highest value of outliers, for instance, if two outliers are found in step (b), assume the case of three outliers while maintaining original values of the mean, standard deviation and number of measurements.

(f) Repeat steps (b) through (d) sequentially increasing the number of outlier possibilities until no more data measurements needs elimination.

(g) Obtain the new value for the mean and sample standard deviation of the reduced data set.

It should be noted that Peirce's Criterion assumes that the data set emanates from a uniform distribution.

2.2 Chauvenet's criterion

One method that has gained wide acceptance in the rational identification of outlier is the Chauvenet's criterion. Peirce's criterion has been buried in the scientific literature for approximately 150 years and is virtually unknown today. Taylor [12] proposed that Chauvenet's criterion is a means of assessing whether one or more pieces of outliers is lodged in a set of observations. Like the Peirce's criterion, Chauvenet's criterion also uses the elements of probability theory to develop an identification format for outliers while assuming that the data set is from a Uniform distribution. The criterion states that all data points should be retained that falls within a bound around the mean that corresponds to a probability of 1 - $\frac{1}{2N}$. In other word, data points can be considered for rejection only if the probability of obtaining their deviation from the mean is less than $\frac{1}{2N}$. The Chauvenet's criterion of identifying outliers uses the following procedures

(i) Calculate the mean and standard deviation of the observed data.

(ii) Find the value of $\frac{d}{\sigma}$ from the Chauvenet's table for the corresponding value of N where, d = deviation, N = number of observations, σ = standard deviation of the distribution.

(iii) Calculate the critical deviation $d_c = S\left(\frac{d}{\sigma}\right)$ where, S = estimate of σ

(iv) If the observed deviation (mean – X_i) is greater than the critical deviation d_c, then such observation is identified as an outlier.

It should be noted that only one observation can be identified at a time.

2.3 Grubbs' test

Grubbs' test is used to detect outliers in a univariate data set. It detects outliers one at a time. When the outlier is detected, it is expunged iterated until no outliers are further detected. However, multiple iterations change the probabilities of detection. Thus, the test should not be used for sample sizes of six or less since when this is done it frequently tags most of the points as outliers.

Grubbs' test is also known as the maximum normed residual test and is based on the normality assumption. The Grubbs' test statistic is defined as the largest absolute deviation from the sample mean in units of the sample standard deviation. This is given by G = $\frac{\max|y_i - \bar{y}|}{S}$, where, $\bar{y}$ = sample mean, y_i = sample observation (suspected outlier), S = sample standard deviation. This test statistic is the largest absolute deviation from the sample mean in units of the sample standard deviation. This is the two-sided version of the test. The Grubbs' test can also be defined as one of the following one-sided tests;

(i) test whether the minimum value is an outlier G = $\frac{\bar{y} - y_{\min}}{S}$, where, y_{min} is the minimum value

(ii) test whether the maximum value is an outlier G = $\frac{y_{\max} - \bar{y}}{S}$, where, y_{max} is the maximum value. The Grubbs' test for identifying outliers has the following procedures:

(a) calculate the sample mean and sample standard deviation

(b) calculate the Grubbs' test statistic G

(c) find the value of critical G from the Grubbs' table for the corresponding value of N.

(d) if the value of the Grubbs' test statistic G is greater than the value of the critical G then, such an observation is an outlier.

2.4 Comparison between Chauvenet's and Peirces' criteria

In the cause of this paper, we have taken to give a detailed review of three identification methods but in this aspect, we shall take into consideration two closely knit methods which has been the cause of several heated debates in the statistical world. The two methods are the Peirce's

and the Chauvenet's Criteria.

These two criteria are both based on probability theory and both assume data sets to be from normal distributions Chauvenet's Criterion makes an arbitrary assumption concerning the identification of outlier data while Peirce's Criterion does not make such assumption. Chauvenet's Criterion makes no distinction between one or several suspicious data values while Peirce's Criterion is a rigorous theory that can be easily applied to several outlier data values. Finally, both criteria are in some ways similar to each other, yet in so many other ways different.

3.0 Empirical illustration

Data on the economic data page of www.nigeriabusinessinfo.com [14] was reviewed. The data was on the graduate output by discipline in Nigeria from 1993 to 1997.

(3.1) Identifying outliers using peirce's criterion

To detect outliers in the data, the following steps were used:

(a) The mean and standard deviation are 1,451.4 and 1,758.1 respectively

(b) Let there be one suspect data in the data set of forty six measurements. Then, the *R*-value from Peirce's table is 2.560

(c) To calculate

(i) $|x_i - \text{mean}|_{\max} = R\text{-value } x \text{ standard deviation}$

Maximum allowable deviation = 2.560 x 1,758.1 = 4,500.7

(ii) But $|x_i - \text{mean}|$ for $x_i = 8{,}962$ is 7,510.6

Since $|x_i - \text{mean}| > |x_i - \text{mean}|_{\max}$, the data point is an outlier

(d) Since only one outlier has been identified we now assume two more doubtful observations and apply step (c) again:

(i) To calculate, $|x_i - \text{mean}|_{\max}$ = R-value *x* standard deviation. R-value for two doubtful observations for $N = 46$ is 2.290, $S = 1{,}758.1$. Therefore, $|x_i - \text{mean}|_{\max} = 4{,}026.1$

(ii) but $|x_i - \text{mean}|$ for $x_i = 5{,}818$ is 4,363.6

(iii) for $x_i = 5{,}271$, $|x_i - \text{mean}| = 3{,}819.6$. For the data point 5,818, $|x_i - \text{mean}| > |x_i - \text{mean}|_{\max}$

Thus, the data point 5,818 is an outlier. Meanwhile, using the same procedure, it was observed that the data point 5,271 is not an outlier. It then follows that the Peirce's Criterion has identified the data points 8,962 and 5,818 as the two outliers in the data set. According to Benjamin [3], we delete the outlier points and then calculate the mean and standard deviation of the outlier free data. Hence, the new mean becomes 1,199.2 while the new standard deviation is 1,277.8.

3.2 Identifying outliers using Chauvenet's criterion

Using Chauvenet's criterion, for the identification of outliers, the following steps were used:

(a) The mean and standard deviation are 1,451 and 1,758.1 respectively

(b) For N = 46, using Chauvenet's table, $\frac{d}{\sigma} = 2.53$, by extrapolation

(c) To obtain the critical deviation, we use the following steps

(i) the suspected outlier $x_i = 8{,}962$

(ii) the observed deviation = 7,510.6

(iii) the critical deviation = 4,447.9

(iv) since the observed deviation is greater than the critical deviation, the data point 8,962 is an outlier

(d) To test for another suspected outlier we proceed thus:

(i) suspected outlier $x_i = 5{,}818$

(ii) observed deviation = 4,366.0
(iii) critical deviation = 4,447.9
(iv) since the observed deviation is less than the critical deviation, the data point 5,818 is not an outlier

From this analysis, Chauvenet's criterion has identified only one outlier 8,962 in the data set which

is a contrast to the Peirce's criterion which identified two outliers.

3.3 Identifying outliers using Grubbs' test

The following steps were used in the detection of outliers using Grubbs' Test

(a) The mean and standard deviation are 1,461.4 and 1,758.1 respectively
(b) The suspected outlier is Y_i = 8,962. The Grubbs' test statistic G is 4.27
(c) From Grubbs' table, Critical Z for N = 46 is 3.09
(d) Since the Grubbs' test statistic G is greater than the critical Z, the data point 8,962 is an outlier. Hence, to test for another suspected outlier Y_i = 5818. The Grubbs' test statistic G = 2.48

Comparing this with the critical $\cancel{Z}$, the data point 5,818 is not an outlier
The Grubbs' test like the Chauvenet's criterion identified only one outlier from the data set.

4.0 Summary

In summary, identification of outliers allows for proper dealings with outliers which in turn gives better and unbiased outcomes to the analysis. It should also be noted that in using any of these three methods of identifying outliers in a data set, before a data set is outlier free, both the minimum and maximum data points must be confirmed as non outliers. When an outlier is identified and properly dealt with, accuracy tends to increase significantly in data analysis. This is the main reason why many researchers lay emphasis on identifying outliers firstly before any other step in any data analysis. The summary of the analysis is given in the Table 4.1:

Table 4.1: Tabulated analysis using the three methods of identifying outliers.

	Peirce's Criterion	**Chauvenet's Criterion**	**Grubbs'Test**
Assumptions	1) Data is normally distributed. 2) Presence of more than one outlier at a time.	1) Data is normally distributed. 2) Presence of one outlier at a time.	1) Data is normally distributed. 2) Presence of one outlier at a time.
Parameters needed	Mean, standard deviation, R-value from the Peirce's table.	Mean, standard deviation, $\frac{d}{\sigma}$ value from Chauvenet's table.	Mean, standard deviation, $\frac{d}{\sigma}$ value from Grubbs' table.
Result	Two outliers, 8,962 and 5,818	One outlier, 8,962	One outlier, 8,962
Conclusion	Delete the outliers	Delete the outlier	Delete the outlier.

References

[1] Anscombe, F.T. (1960). Rejection of outliers. Technometrics, 2, 123 – 147.
[2] Barnett, V. and Lewis, T. (1994). Outleirs in statistical data (3rd ed.) New York: Wiley.
[3] Benjamin, P. (1852). Criterion for rejection of doubtful observations, Astronomical journal II, 45, (1854).
[4] Fox, A.J. (1972). Outliers in time series. J. Roy. Stat. Soc. 13, 43, 350 – 363.
[5] Hamilton, L.C. (1992). Regressions with graphics: a second course in applied statistics. Monterey, C.A.: Brooks and Cole.
[6] Harvey, M. (2000). Outliers make statistical analysis difficult. Jour. of Stat. Analysis. 10. 321 – 330.
[7] Hawkins, D.M. (1980). Identification of outliers. London Chapman and Hall.
[8] Judd, C.M. and McClelland, G.H. (1989). Data analysis, a model comparison approach. San Diego, CA: Harcourt Brace Jovanovich.
[9] Osborne, J.W. and Overbay, A. (2004). The power of outliers (and why researchers should always check for them). Practical assessment, research and evaluation. 916.

[10] Rasmussen, J.L. (1988). Evaluating outlier identification tests: Mahalanobis D squared and Comrey D Multivariate behavioral research, 23 (2), 189-202.
[11] Ross, S. (2003). University of New Haven article, J. Eng. Technology.
[12] Taylor, J.R. (1997). An introduction to error analysis, 2nd ed. Sausolito, CA. Univ. Science books, pp 166 – 168.
[13] Tsay, R.S. (1988). Outliers, level shifts and variance changes in time series. J. forecasting, 7, 1 – 20.
[14] Wikipedia Website (www.wikipedia.org/outleirs).