Why don't simple outlier methods work? The argument against our competition.

Posted on Aug 01, Posted by [Tom Reilly](#) Category [Forecasting](#)

Why don't simple outlier methods work? The argument against our competition.

For a couple of reasons:

It wasn't an outlier. It was a seasonal pulse.

The observations outside of the 2 or 3 sigma bounds could in fact be a newly formed seasonal pattern. For example, halfway through the time series June's become become very high when it had been average. Simple approaches would just remove anything outside the bounds which could be throwing the "baby out with the bathwater".

Your 3 sigma calculation was skewed due to the outlier itself.

It is a chicken and egg dilemma. The outliers make the sigma wide so that you miss outliers.

The outlier was in fact a promotion.

Using just the history of the series is not enough. You should include causals as they can help explain what is perceived to be an outlier.

Now let's consider the inlier.

There could be outliers that are within 3 sigma and let's say the observation is near the mean. When could the mean be unusual? When the observation should have been high and it just didn't for some reason.

Simple methods force the user to specify the # of times the system should iterate to remove outliers.

You are then asked how many times do you want to iterate to find the interventions by the forecasting tool? Is this intelligence or a crutch? So, you are somehow supposed to provide some empirically based guidance??? You don't know as it would be just a guess.

The reality is that Simple methods/software use a process where they assume a "mean model" to determine the outliers. The correct way is to build a model and identify the outliers at the same time. Sounds simple, right?

Refer to these articles for more on how to identify outliers properly

Fox JA (1972). Outliers in time series. J. Royal Stat. Soc., Series B, 34: 350-363.

Chang, I., and Tiao, G.C. (1983). "Estimation of Time Series Parameters in the Presence of Outliers," Technical Report #8, Statistics Research Center, Graduate School of Business, University of Chicago, Chicago.

Tsay R (1986a). Time series model specification in the presence of outliers. J. Am. Stat. Soc., 81: 132-141.

Tsay R (1988). Outliers, level shifts and variance changes in time series. J. Forecast., 7: 1-20.

Does anyone have any other examples of bad outlier methodologies? or other software with their examples posted?

Tags: Tagged in: [outliers fox tiao tsay box box-jenkins](#)