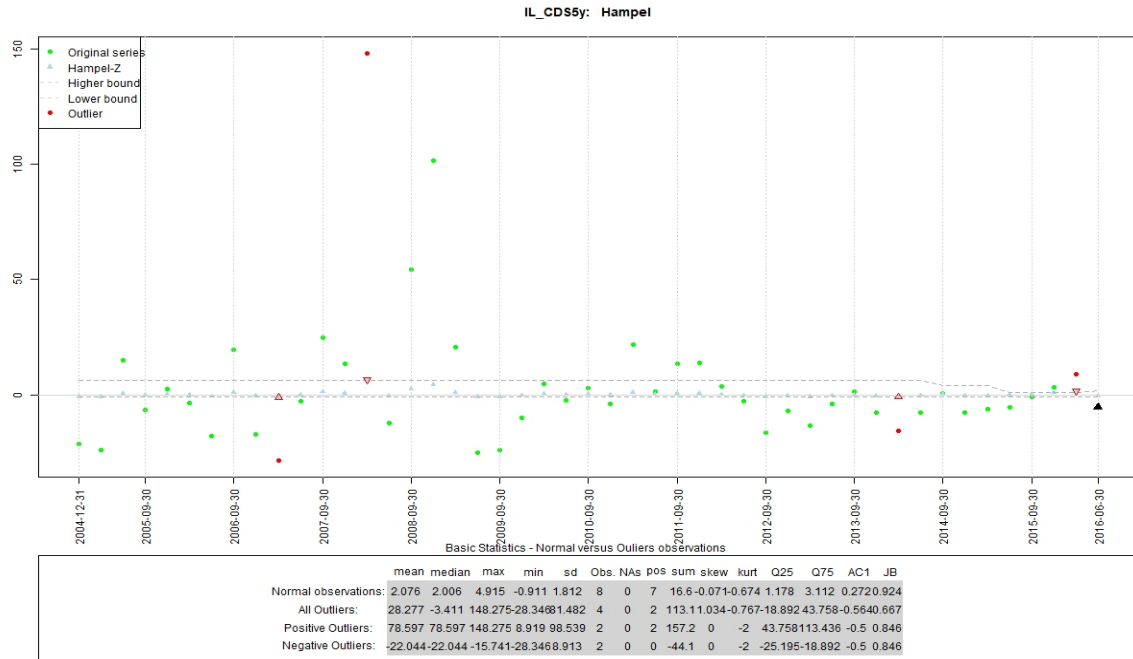


Help on Hampel (1971) method

Hampel's method is a robust version of outliers detection as it is based on medians rather than means. It is based on robust estimators for the location and the scale of the distribution, i.e. median and MAD (Median Absolute Deviations) rather than the mean and standard deviation. Hampel's limits are:

$$Limit = Median(x) \pm z_{1-\alpha/2}MAD(x)$$

where $MAD = 1.4826 \cdot Median(|x - Median(x)|)$ is a scale estimator and $z_{1-\alpha/2}$ is the $1-\alpha/2$ percentile of the z distribution. For a normally distributed sample, MAD is a consistent estimator of the standard deviation, and α is the type 1 error for detecting outliers. For each observation in the series, x , an $Hampel_Z = \frac{x - Median(x)}{MAD_x}$ is calculated and if that mark is higher (lower) than the upper (lower) limit it is considered an outlier. Screen example with high and low percentiles of 0.995 and 0.005 (i.e., $\alpha/2 = 0.995$ and $1 - \alpha/2 = 0.005$), respectively based on the 100 last observations (weeks), is depicted below. Note that the limits (gray broken lines) are not constant. For instance, after a period (100 weeks in our case) of high volatility MAD grows, thus the limits become narrower. Gray triangles represent the $Hampel_Z$ of the respective normal observations while red circled triangles point to outliers (in red). Particularly, upward (red circled) triangles point to respective negative outliers while downward ones point to respective positive outliers. A summary statistics of both normal observations, positive outliers, negative outliers, and all outliers are exhibited at the bottom of the figure.



References

Hampel, R.F. (1971), “A General Qualitative Definition of Robustness,” *The Annals of Mathematical Statistics*, 42, 1887–1896.