An analysis using simulation to compare several moving average techniques for time series data.

A Preprint

MD SAKIBUR HASAN
Department of Statistics
Grand Valley State University
Campus Dr, Allendale, MI 49401, United States
shstat10@gmail.com

Bishal sarker
Department of Applied Statistics
University of Dhaka
Dhaka, Bangladesh
sarkerbishal02@gmail.com

January 30, 2023

Abstract

There are many smoothing techniques available, and selecting the appropriate Technique is a very important issue to achieve good smoothing performance. This study intends to compare different types of moving average smoothing techniques. The study compares the simple moving average (SMA), exponential moving average (EMA), double exponential moving average (DEMA), weighted moving average (WMA), and zero-lag exponential moving average (ZLEMA).Some error measures are mean error, mean absolute error, and mean square error. are calculated for the above smoothing techniques to compare the smoothing accuracy of these methods.The simulated data used for the study came from six different statistical distributions.The study assists in determining the best moving average smoothing method for any time series data.

***K*eywords** SMA · WMA · ZLEMA · DEMP · Moving average · EMA

1 Introduction:

1.1 Background of The Study

Data for time series is gathered from various points throughout time. As a result, the data set has a great deal of variety. So, a technique known as smoothing is employed to lessen these variations. Techniques for removing noise from a time series of data include smoothing techniques. It aids in determining the datset's trend. When data is compiled, any volatility or other types of noise can be removed or reduced. Data smoothing is the term for this. Data smoothing is based on the notion that it can recognize simpler changes to assist in the prediction of various trends and patterns. It serves as a tool for statisticians or traders who must examine a lot of data, which is frequently challenging.

1.2 Objective of The Study

Among all others smoothing methods moving average methods is the oldest and simplest smoothing methods. The main object of that study is comparing different types of moving average smoothing techniques such as simple moving average(SMA),exponentially-weighted moving average (EWMA), weighted moving average (WMA), double exponential moving average(DEMA),Hull moving average(HMA), Zero lag exponential moving average(ZLEMA) etc. Among them the syudy compares SMA, EWMA, WMA, DEMA,HMA and ZLEMA.

2 Literature review

Raudys, Lenčiauskas, and Malčius (2013) smoothed financial data using the moving average.

Ivanovski, Milenkovski, and Narasanov (2018) extrapolated the number of tourists using the moving average. The study aids in making wise decisions for the future.

Hameed (2015) compares every smoothing method currently in use to forecast future demand for private universities in Bangladesh. They contrast different kinds of currently used smoothing techniques and discover that Holt's method provides the optimum accuracy for their work.

For the purpose of early detection of infectious disease outbreaks, Yang et al. (2018) conducts simulation-based studies on the comparison of statistical and time series forecasting techniques. Here, various approaches are discussed in an effort to use simulation to produce the greatest results.

Sinaga and Irawati (2020) studied about medical disposable supply demand forecasting by moving average and exponential moving average method.

For predicting power load, Karim and Alwi (2013) employed exponential smoothing and moving average, and exponential moving average outperformed moving average.

Fong et al. (2020) tried to find an accurate early forecasting model from Small dataset of 2019-nCoV Novel Coronavirus outbreak.

3 Methodology

3.1 Simple moving average (SMA)

By averaging several different subsets of the entire data set, a simple moving average (SMA) statistical technique is utilized to examine data points. Because it is calculated by averaging a predetermined number of successive data points with equal weight for each, it is referred to as being "simple." The SMA is used to highlight long-term trends or patterns in the data and to smooth out short-term volatility in the data.The formula for SMA is given below,

$$T_t = \frac{\sum_{i=-m}^{m} Y_{t+i}}{k}$$

here k is the number of order and Y_i is the observation. m=$\frac{k-1}{2}$ is the half width of a moving average as the number of points. For example, a arithmetic moving average of 3 ordered at time t is $T_t = \frac{Y_{t-1}+Y_t+Y_{t+1}}{3}$.

Simple Moving Average (SMA) and Exponential Moving Average (EMA) both measure trend direction over time. SMA only determines an average, but EMA gives more weight to data that is more recent.

3.2 Exponentially moving average (EMA)

Simple Moving Average (SMA) and Exponential Moving Average (EMA) both measure trend direction over time. SMA only determines an average, but EMA gives more weight to data that is more recent.

$$EMA = C\breve{\ }P\frac{2}{(n+1)} + P$$

where C and P are current data point and an exponential moving average of the previous period (simple average used for the first period) respectively. THe formula is given below,

3.3 Weighted moving average

When calculating the weighted moving average, recent data points are given more weighting than historical data points. When added together, the weights' total value ought to be 100%, or 1. The weighting factor used to calculate the WMA is determined by the period selected for the indicator. For example, a 5 period WMA would be calculated as follows:

$$WMA = \frac{(5P_1 + 4P_2 + 3P_3 + 2P_4 + 1P_5)}{(5+4+3+2+1)}$$

Where, P_1= current price P_2 = price one bar ago and so on.

3.4 ZLEMA

John Ehlers and Ric Way created the Zero Lag Exponential Moving Average (ZLEMA) indicator. The goal is to get rid of the inherent lag that all averages and other trend following indicators have.

This is what the ZLEMA aims to accomplish by tracking recent prices more closely than historical prices, much like a standard EMA but with an even greater emphasis on recent prices. A moving average with less lag and good smoothing is the end product.

The benefits of using zero lag moving averages are as follows:

- A moving average exponential without lag that is more responsive to current price changes.
- The indicator is applicable across all instruments and timeframes.
- The indicator may be used as a signal or as a filter for signals.

Formula are given below ,

$$\alpha = \frac{2}{n+1}$$

$$Z_t = (1-\alpha)Z_t + \alpha(T_t - T_{t-\frac{n-1}{2}})$$

Where , n is the number of period. α represents the lag and Z_t represents the ZLEMA moving average at t time points.

3.5 DEMA

A level component and a trend component are used in double exponential smoothing at each period. Two weights, also known as smoothing parameters, are used in double exponential smoothing to update the components at each time.The formula is given below,

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \gamma(L_t{}^{\smile}L_{t^{\smile}1}) + (1-\gamma)T_{t^{\smile}1}$$

$$\hat{Y} = L_{t^{\smile}1} + T_{t^{\smile}1}$$

Here , L_t level at time t.
α weight for the level.
T_t trend at time t
γ weight for the trend
Y_t data value at time t
$\hat{Y}_t$ fitted value, or one-step-ahead forecast, at time t

3.6 Mean absolute error (MAE)

Mean absolute error (MAE), which measures errors between fitted and observed values, is a statistical concept. Comparisons of expected against observed data, subsequent time against initial time, and one measuring technique against an alternate measurement technique are a few examples of Y vs X. The MAE is calculated by dividing the total absolute errors by the sample size.

$$MAE = \frac{\sum_{i=1}^{n} |Y_i - \hat{Y}_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}$$

Here $\hat{Y}_i$ is the predicted value and Y_i is the true value.

3.7 Mean square error (MSE)

The mean square error (MSE), which is derived as the total of the squared discrepancies between the forecasts and actual values, divided by the number of data points, is a metric for the average difference between actual and predicted values in a dataset.

$$MSE = \frac{\sum_i^n (\hat{Y}i - Yi)^2}{n}$$

Here $\hat{Y}_i$ is the prediction and Y_i is the true value.

4 Analysis

Here 100000 values are generated from six different distribution. Data are generated from Normal distribution with N $(\mu = 50, \sigma = 5)$, Poisson P $(\lambda = 10)$, Gamma distribution $Gamma$ $(\alpha = 10, \beta = 2)$, T distribution T_{10}, exponential distribution $exp(\lambda = 5)$ and weibull distribution $weib$ $(\alpha = 10, \beta = 5)$. After generating the data, transform it as a time series data. A comparison table among all the moving average techniques used for that study is given below.

Table 1: Smoothing comparison table

Normal Distribution			Poisson Distribution			Weibull Distribution		
TYPE	MAE	MSE	TYPE	MAE	MSE	TYPE	MAE	MSE
SMA	2.830	12.579	SMA	1.776	5.033	SMA	0.320	0.165
EMA	1.634	4.192	EMA	1.030	1.675	EMA	0.185	0.055
DEMA	0.746	0.873	DEMA	0.470	0.349	DEMA	0.084	0.011
WMA	1.887	5.591	WMA	1.184	2.237	WMA	0.213	0.073
ZLEMA	1.885	5.585	ZLEMA	1.187	2.224	ZLEMA	0.214	0.073

4.1 Smoothing comparison table for time period 2 and 100000 values

Table 2: Smoothing comparison table

Gamma Distribution			T Distribution			Exponential Distribution		
TYPE	MAE	MSE	TYPE	MAE	MSE	TYPE	MAE	MSE
SMA	1.776	5.033	SMA	0.619	0.626	SMA	0.100	0.020
EMA	1.030	1.675	EMA	0.358	0.209	EMA	0.059	0.007
DEMA	0.470	0.349	DEMA	0.163	0.043	DEMA	0.027	0.001
WMA	1.184	2.237	WMA	0.413	0.278	WMA	0.067	0.009
ZLEMA	1.187	2.224	ZLEMA	0.415	0.278	ZLEMA	0.069	0.009

4.2 Smoothing comparison table for 5 time period and 10000 values.

Table 3: Smoothing comparison table

Normal Distribution			Poisson Distribution			Weibull Distribution		
TYPE	MAE	MSE	TYPE	MAE	MSE	TYPE	MAE	MSE
SMA	3.562	19.830	SMA	2.250	8.039	SMA	4.689	34.957
EMA	2.912	13.270	EMA	1.845	5.365	EMA	3.584	20.613
DEMA	2.233	1.011	DEMA	1.413	3.144	DEMA	2.735	12.083
WMA	3.036	14.391	WMA	1.916	5.805	WMA	3.723	22.277
ZLEMA	2.617	10.742	ZLEMA	1.652	4.275	ZLEMA	3.197	16.354

Table 4: Smoothing comparison table

Gamma Distribution			T Distribution			Exponential Distribution		
TYPE	MAE	MSE	TYPE	MAE	MSE	TYPE	MAE	MSE
SMA	0.786	1.001	SMA	0.129	0.031	SMA	0.401	0.256
EMA	0.640	0.665	EMA	0.105	0.020	EMA	0.328	0.171
DEMA	0.492	0.390	DEMA	0.080	0.012	DEMA	0.251	0.101
WMA	0.667	0.720	WMA	0.109	0.022	WMA	0.341	0.185
ZLEMA	0.577	0.534	ZLEMA	0.096	0.016	ZLEMA	0.293	0.137

4.3 Smoothing comparison table for 7 time period and 1000 values.

Table 5: Smoothing comparison table

Normal Distribution			Poisson Distribution			Weibull Distribution		
TYPE	MAE	MSE	TYPE	MAE	MSE	TYPE	MAE	MSE
SMA	3.600	20.884	SMA	2.363	8.665	SMA	4.739	35.694
EMA	3.110	15.570	EMA	2.025	6.412	EMA	3.997	25.331
DEMA	2.576	-27833.008	DEMA	1.659	4.379	DEMA	3.307	17.418
WMA	3.202	16.454	WMA	2.081	6.783	WMA	4.107	26.764
ZLEMA	2.826	12.946	ZLEMA	1.813	5.273	ZLEMA	3.640	21.278

Table 6: Smoothing comparison table

Gamma Distribution			T Distribution			Exponential Distribution		
TYPE	MAE	MSE	TYPE	MAE	MSE	TYPE	MAE	MSE
SMA	0.823	1.111	SMA	0.130	0.032	SMA	0.425	0.275
EMA	0.717	0.844	EMA	0.113	0.024	EMA	0.364	0.205
DEMA	0.604	0.585	DEMA	0.093	0.015	DEMA	0.299	0.137
WMA	0.740	0.895	WMA	0.116	0.025	WMA	0.374	0.216
ZLEMA	0.676	0.720	ZLEMA	0.105	0.020	ZLEMA	0.330	0.169

5 Findings and conclusion

From the smoothing comparison table, it is seen that for any datasets and any time period, DEMA has the lowest mean square error among all the moving average smoothing techniques. After DEMA, EMA has the lowest mean square error. Among all the moving average smoothing techniques used for the study, SMA has the highest MSE value. SMA is the simplest and most straightforward method for smoothing, but it performs the worst. DEMA has complex smoothing techniques, but it performs far better than other techniques.

Reference

Fong, Simon James, Gloria Li, Nilanjan Dey, Rubén González Crespo, and Enrique Herrera-Viedma. 2020. "Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-Ncov Novel Coronavirus Outbreak." *arXiv Preprint arXiv:2003.10776.*

Hameed, Haifaa Hussein. 2015. "Smoothing Techniques for Time Series Forecasting." Master's thesis, Eastern Mediterranean University (EMU)-Doğu Akdeniz Üniversitesi (DAÜ).

Ivanovski, Zoran, Ace Milenkovski, and Zoran Narasanov. 2018. "TIME SERIES FORECASTING USING a MOVING AVERAGE MODEL FOR EXTRAPOLATION OF NUMBER OF TOURIST." *UTMS Journal of Economics* 9 (2).

Karim, SA Abdul, and Saiful Azli Alwi. 2013. "Electricity Load Forecasting in UTP Using Moving Averages and Exponential Smoothing Techniques." *Applied Mathematical Sciences* 7 (77-80): 4003–14.

Raudys, Aistis, Vaidotas Lenčiauskas, and Edmundas Malčius. 2013. "Moving Averages for Financial Data Smoothing." In *Information and Software Technologies: 19th International Conference, ICIST 2013, Kaunas, Lithuania, October 2013. Proceedings 19*, 34–45. Springer.

Sinaga, H, and N Irawati. 2020. "A Medical Disposable Supply Demand Forecasting by Moving Average and Exponential Smoothing Method." In *Proceedings of the 2nd Workshop on Multidisciplinary and Applications (WMA) 2018, 24-25 January 2018, Padang, Indonesia.*

Yang, Eunjoo, Hyun Woo Park, Yeon Hwa Choi, Jusim Kim, Lkhagvadorj Munkhdalai, Ibrahim Musa, and Keun Ho Ryu. 2018. "A Simulation-Based Study on the Comparison of Statistical and Time Series Forecasting Methods for Early Detection of Infectious Disease Outbreaks." *International Journal of Environmental Research and Public Health* 15 (5): 966.