Forecast Precision and Forecast Accuracy from Moving Average and Moving Median Methods on Skewed Lognormal Time Series

Louie Ren
University of Houston-Victoria

Peter Ren
University of Houston-Downtown

We examine the forecast precision and accuracy for forecasts from moving average and moving median methods on skewed i.i.d. time series following various lognormal probability distributions. Overall, we recommend the Moving Average method, MA, when forecasting time series that follow lognormal distributions.

INTRODUCTION

In forecasting, "precision" measures the un-biasness of a forecasting method and "accuracy" measures the dispersions of a forecasting method. In absolute terms, the sample mean for the error terms, $\bar{e}$, and the sample standard deviation for the error terms, s_e, are the most popular measures for forecast precision and accuracy, respectively. The inverse coefficient of variation, $\frac{\bar{e}}{s_e}$ (Hoover, 2006), and the coefficient of variation, $\frac{s_e}{\bar{e}}$ (Anderson, Sweeney, Williams, Camm, and Cochran 2016), are the two most popular scale-free measures for forecast precision and accuracy, respectively.

In this study, we introduce a new forecast measure from the linear combination of scale-free forecast measures:

$$L_\omega = \omega\left(\frac{\bar{e}}{s_e}\right) + (1-\omega)\left(\frac{s_e}{\bar{e}}\right), \text{ where } 0 \leq \omega \leq 1.$$

The weight, ω, serves to reflect the forecaster's emphasis on precision relative to accuracy. When $\omega = 0$, $L_{\omega=0}$ is the popular scale free forecast accuracy measure, the coefficient of variation, $\frac{s_e}{\bar{e}}$. When $\omega = 1$, $L_{\omega=1}$ is the popular scale free forecast accuracy measure, the inverse coefficient of variation, $\frac{\bar{e}}{s_e}$. For instance, when $\omega=0.2$, the forecaster is more concerned (80%) about the forecast accuracy from $\frac{s_e}{\bar{e}}$ rather than the forecast precision from $\frac{\bar{e}}{s_e}$.

In this study, we compare forecast precisions and accuracies from the forecasting methods Moving Average (MA) and Moving Median (MMd) on various skewed i.i.d. time series following lognormal distributions. Forecast precisions and accuracies are measured by absolute measures, $\bar{e}$ and s_e, and relative measures L_ω for ω = 0, 0.2, 0.4, 0.6, 0.8, and, 1.0.

We find that out of skewed i.i.d. time series from lognormal distributions with parameters (μ =0, σ =1), (μ =0, σ =0.5), (μ =0, σ =0.25), (μ =0, σ =0.1) and skewness $(e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$, 6.1849, 1.9446, 0.8340, 0.3113, respectively, the Moving Median method, MMd, only outperforms the Moving Average method, MA, in terms of forecasting accuracy for extremely skewed lognormal distribution with parameters (μ =0, σ =1), and skewness 6.1849. Overall, we recommend the Moving Average method, MA, when forecasting time series that follow lognormal distributions.

DATA ANALYSIS

In this study, we compare forecast precisions and accuracies from the forecasting methods Moving Average (MA) and Moving Median (MMd) on various skewed i.i.d. time series following lognormal distributions. Forecast precisions and accuracies are measured by absolute measures, $\bar{e}$ and s_e, and relative measures L_α for α = 0, 0.2, 0.4, 0.6, 0.8, and, 1.0.

A lognormal probability distribution is a probability distribution for a continuous random variable X > 0, where ln(X) follows a normal probability distribution. In summary, when X follows a lognormal distribution and ln(X) follows a normal distribution with a mean of μ and a s.d. of σ, its probability density function for X is as follows

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \ x > 0,$$

with E(X) = $e^{(\mu+\frac{\sigma^2}{2})}$, Median(X) = e^{μ}, Var(X) = $(e^{\sigma^2}\text{-}1)e^{(2\mu+\sigma^2)}$, and the Skewness = $(e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$.

Examples of lognormal distributions can be found in the fields of finance, and economics (Johnson, Kotz, and Balakrishnan, 1994, PP. 210-211, and Antoniou, Ivanov, Ivanov, and Zrelov, 2004).
In this study, a simulation of 1,000 observations from a lognormal distribution are grouped into 50 groups each with 20 observations. For each group, the first 10 observations are treated as historical data, and the next 10 observations are treated as the realizations of actual observations. Moving averages (MA) and moving medians (MMd) with moving periods of 10 are applied as our forecasts for period 11 to 20. Forecast errors are defined as the differences between the realizations of actual observations for period 11 to 20 and the forecasts generated from MA and MMd with moving periods of 10. The MA and MMd with moving periods of 10 are denoted as MA(10) and MMd(10), respectively.

Pairwise T-test (or the Student T-test) for testing H_0: $\mu_{MA} \geq \mu_{MMd}$ vs H_a: $\mu_{MA} < \mu_{MMd}$ is applied to the 50 independent $\bar{e}$, s_e, and L_ω for α = 0 (for $\frac{s_e}{\bar{e}}$), 0.2, 0.4, 0.6, 0.8, and 1.0 (for $\frac{\bar{e}}{s_e}$) generated from MA(10) and MMd(10), respectively. The results for the T-test are listed on the following page.

CONCLUSION

In this study, we find that the Moving Median method, MMd, only outperforms the Moving Average method, MA, in terms of forecasting accuracy for time series follow a very skewed lognormal probability distribution with parameters (μ =0, σ =1), and skewness 6.1849. Overall, we recommend the Moving Average method, MA, when forecasting time series that follow lognormal distributions (e.g. stock prices).

TABLE 1

TESTING H_0: $\mu_{e,MMd} \geq \mu_{e,MA}$ vs H_a: $\mu_{e,MMd} < \mu_{e,MA}$ for LOGNORMAL with μ=0, σ=1

	$\overline{e_1} - \overline{e_2}$	$s_{e_1} - s_{e_2}$	$L_{\omega=0}$ $\frac{s_{e_1}}{\overline{e_1}} - \frac{s_{e_2}}{\overline{e_2}}$	$L_{\omega=0.2}$	$L_{\omega=0.4}$	$L_{\omega=0.6}$	$L_{\omega=0.8}$	$L_{\omega=1.0}$ $\frac{\overline{e_1}}{s_{e_1}} - \frac{\overline{e_2}}{s_{e_2}}$
Mean	0.4811	-0.0693	2.5444	2.1086	1.6727	1.2369	0.8010	0.3652
s.d.	0.3442	0.1182	34.0192	27.2158	20.4128	13.6106	6.8116	0.3643
T-ratio	9.8833	-4.1435	0.5289	0.5478	0.5794	0.6426	0.8315	7.0884
p-value	1.0000	0.0001*	0.7004	0.7069	0.7175	0.7383	0.7951	1.0000

* significant at α = 0.01, the skewness = $(e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$ = (e+2)$\sqrt{e - 1}$ ≈ 6.1849

TABLE 2

TESTING H_0: $\mu_{e,MMd} \geq \mu_{e,MA}$ vs H_a: $\mu_{e,MMd} < \mu_{e,MA}$ for LOGNORMAL with μ=0, σ=0.5

	$\overline{e_1} - \overline{e_2}$	$s_{e_1} - s_{e_2}$	$L_{\omega=0}$ $\frac{s_{e_1}}{\overline{e_1}} - \frac{s_{e_2}}{\overline{e_2}}$	$L_{\omega=0.2}$	$L_{\omega=0.4}$	$L_{\omega=0.6}$	$L_{\omega=0.8}$	$L_{\omega=1.0}$ $\frac{\overline{e_1}}{s_{e_1}} - \frac{\overline{e_2}}{s_{e_2}}$
Mean	0.1130	0.0050	6.7375	5.4273	4.1171	2.8069	1.4967	0.1865
s.d.	0.0953	0.0456	41.4876	33.1988	24.9101	16.6214	8.3332	0.1572
T-ratio	8.3767	0.7757	1.1483	1.1560	1.1687	1.1941	1.2700	8.3866
p-value	1.0000	0.7792	0.8718	0.8733	0.8759	0.8809	0.8950	1.0000

* significant at α = 0.01, the skewness ≈ 1.9446

TABLE 3

TESTING H_0: $\mu_{e,MMd} \geq \mu_{e,MA}$ vs H_a: $\mu_{e,MMd} < \mu_{e,MA}$ for LOGNORMAL with μ=0, σ=0.25

	$\overline{e_1} - \overline{e_2}$	$s_{e_1} - s_{e_2}$	$L_{\omega=0}$ $\frac{s_{e_1}}{\overline{e_1}} - \frac{s_{e_2}}{\overline{e_2}}$	$L_{\omega=0.2}$	$L_{\omega=0.4}$	$L_{\omega=0.6}$	$L_{\omega=0.8}$	$L_{\omega=1.0}$ $\frac{\overline{e_1}}{s_{e_1}} - \frac{\overline{e_2}}{s_{e_2}}$
Mean	0.0259	-0.0010	-0.7038	-0.5451	-0.3865	-0.2278	-0.0691	0.0896
s.d.	0.0384	0.0167	23.3133	18.6559	13.9986	9.3414	4.6849	0.1397
T-ratio	4.7810	-0.4333	-0.2135	-0.2066	-0.1952	-0.1724	-0.1043	4.5343
p-value	1.0000	0.3333	0.4159	0.4186	0.4230	0.4319	0.4587	1.0000

* significant at α = 0.01, the skewness ≈ 0.8340

TABLE 4

TESTING H_0: $\mu_{e,MMd} \geq \mu_{e,MA}$ vs H_a: $\mu_{e,MMd} < \mu_{e,MA}$ for Lognormal with μ=0, σ=0.1

	$\overline{e_1} - \overline{e_2}$	$s_{e_1} - s_{e_2}$	$L_{\omega=0}$ $\frac{s_{e_1}}{\overline{e_1}} - \frac{s_{e_2}}{\overline{e_2}}$	$L_{\omega=0.2}$	$L_{\omega=0.4}$	$L_{\omega=0.6}$	$L_{\omega=0.8}$	$L_{\omega=1.0}$ $\frac{\overline{e_1}}{s_{e_1}} - \frac{\overline{e_2}}{s_{e_2}}$
Mean	0.0048	0.0002	-59.8470	-47.866	-35.886	-23.906	-11.926	0.0543
s.d.	0.0152	0.0049	316.020	252.817	189.613	126.410	63.2067	0.1723
T-ratio	2.2445	0.2284	-1.3391	-1.3388	-1.3383	-1.3373	-1.3342	2.2282
p-value	0.9853	0.5899	0.0934	0.0934	0.0935	0.0937	0.0942	0.9848

* significant at α = 0.01, the skewness ≈ 0.3113

REFERENCES

Anderson, D., Sweeney, D., Williams, T., Camm, J., & Cochran, J. (2016). *Modern Business Statistics with Microsoft Office Excel*. 6th ed., Boston, MA: Cengage Learning.

Antoniou, I., Ivanov, Vi., Ivanov Va., & Zrelov, P. (2004). On the log-normal distribution of stock market data. *Physica A*, 331, 617-638.

Armstrong, J., & Collopy, F. (1992). Error measure for generalizing about forecasting methods: empirical comparisons. *International Journal of Forecasting,* 8, 69-80.

Baldwin, M. E., & Kain, J. S. (2006). Sensitivity of several performance measures to displacement error, bias, and event frequency. *Weather & Forecasting*, 21(4), 636-648.

Chatfield, C. (1998). Apples, oranges, and mean squared error. *International Journal of Forecasting*, 4, 515-518.

Chen, C., Twycross, J., & Garibaldi, J. M. (2017). A new accuracy measure based on bounded relative error for time series forecasting. *Plos ONE*, 12(3), 1-23.

Flores, B. E., & Wichern, D. W. (2005). Evaluating forecasts: a look at aggregate bias and accuracy measures. *Journal of Forecasting*, 24(6), 433-451.

Hoover, J. (2006), Measuring forecast accuracy: omissions in today's forecasting engines and demand-planning software. *Foresight*, 4, 32-35.

Johnson, N., Kotz, S., & Balakrishnan, N. (1997). *Continuous Univariate Distributions*, Ch. 14(1), 2nd ed., New York, New York: John Wiley & Sons.

Limpert, E., Stahel, W, & Abbt M. (2001). Lognormal distributions across the sciences: keys and clues. *BioScience*, 51(5), 341-352.

Makridakis S. (1993). Accuracy measures: theoretic and practical concerns. *International Journal of Forecasting*, 9, 527-529.

Mazais, J., Miķelsons, J., & Salenieks, N. (2007). Process parameter and measurement accuracy and precision. *Transport & Engineering*, 2311-15.

Morlidge, S. (2016). Using error analysis to improve forecast performance. Foresight: *International Journal of Applied Forecasting*, 41, 37-44.

Shang, H. L. (2015). Statistically tested comparisons of the accuracy of forecasting methods for age-specific and sex-specific mortality and life expectancy. *Population Studies*, 69(3), 317-335.