STAT 153 Project, Analysis on Price of Lots-of-stuff Incorporated Stock

Junlin Liu

February 11, 2021

1 Executive Summary

Lots-of-stuff Incorporated's stock (LOSI) price is expected to be stable for the first two weeks of 2021. Following the rapid changing of market of the year of 2020 because of the COVID-19 pandemic, the stock market tends to be stable by the end of the year of 2020. For the first two weeks of 2021, LOSI would remain at a stable level.

2 Exploratory Data Analysis

The LOSI stock price has been growing annually, as seen below in Figure @ref(fig:EDA). There is no strong seasonality pattern shown, but the several issues impacted the stock price. Through out 2018 and early-2019, the economic crisis impacted the company's stock price, but soon recovered in late-2019. In March 2020, the stock market crash also hhad huge impact on the company's stock price, but soon recovered as well.

3 Models Considered

To model the natural signal in this data, both a parametric model and a differencing approach are used. Both of these models of the signal will be complimented with ARMA models for the remaining noise.

3.1 Parametric Signal Model

First, a parametric model is considered. But seasonality does not seem to exist. An indicator for the impact during 2018 and early-2019 is added, and another indicator for the the March 2020 stock market crashh caused by the outbreak of COVID-19 is also added. This deterministic signal model is detailed in Equation below, where X_t is the additive noise term.

$$\text{Price}_t = \beta_0 + \beta_1 t + \beta_2 I_{\text{eighteen}_t} + \beta_3 t I_{\text{eighteen}_t} + \beta_4 I_{\text{covid}_t} + X_t$$

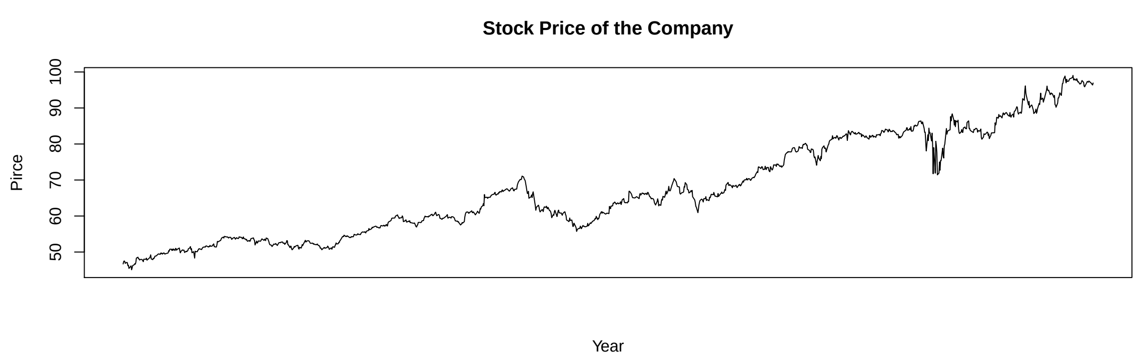

Figure 1: Stock price of the company.

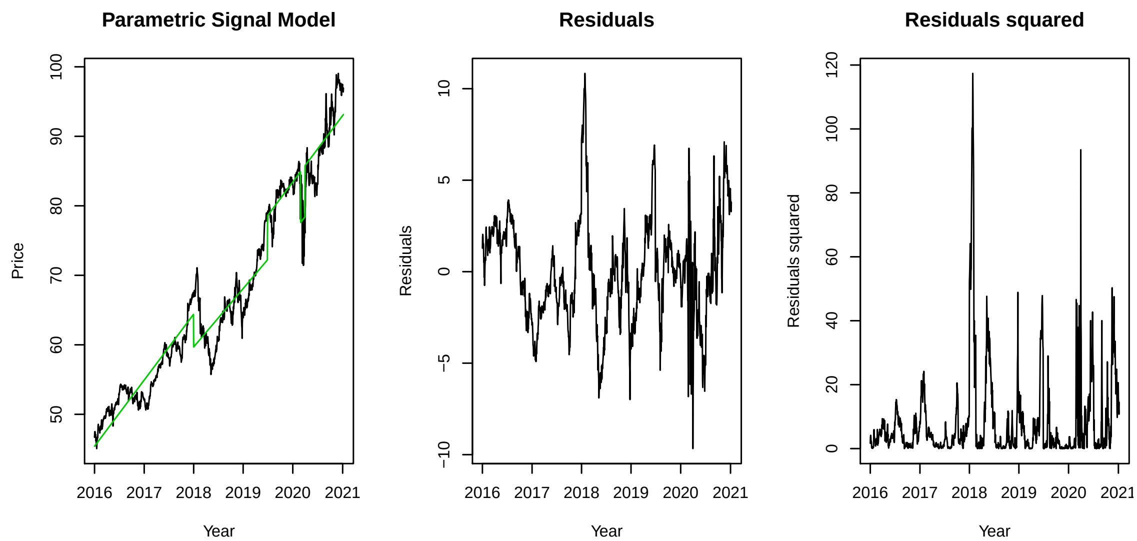

Figure 2: The parametric signal model. The left panel shows this model's fitted values in green, plotted atop the stock data in black. The middle panel shows the residuals of this model. The right panel shows the residual squares of this model.

Figure @ref(fig:signal1) presents the fit as well as the residuals, which appear reasonably stationary. Both plots focus in on the last two years of data in order to show the fine details.

3.1.1 Parametric Signal with AR(10)

The ACF and PACF plots for the parametric model residuals are shown in Figure @ref(fig:acf1). ACF shows that MA models can not fit in this case. The lags with the largest magnitude PACF values occur at lags 1, 2, 9, and 10. Therefore, AR(10) is proposed, and this model implies the ACF and PACF indicated by the red circles in Figure @ref(fig:acf1) which fit the general pattern of the sample autocorrelations.

3.1.2 Parametric Signal with AR(13)

The ACF and PACF plots for the parametric model residuals are shown in Figure @ref(fig:acf1). ACF shows that MA models can not fit in this case. The PACF also shows relatively significant magitudes at 1, 10, and 23. Therefore, given that AR(10) has been tried AR(13) can be another try, and this model implies the ACF and PACF indicated by the blue circles in Figure @ref(fig:acf1) which fit the general pattern of the sample autocorrelations.

3.2 Differencing

As shown in Figure @ref(fig:EDA), the stock price shows some non-linear trend, whereas the first order differencing does not fit well in this case. Thus, a second order differencing will be helpful. Since there shows no seasonality, a simple differencing would solve the case.

3.2.1 Differencing with MA(1)

The sample ACF and PACF for the second order differences are shown in Figure @ref(fig:acf2). The lags with the largest magnitude ACF values occur at lags 1 and 2. The PACF shows a exponential decease pattern. As a result, MA(1) is a good fit in this case. The red circles show that this MA(1) specification fits the general deseace pattern in the PACF plot.

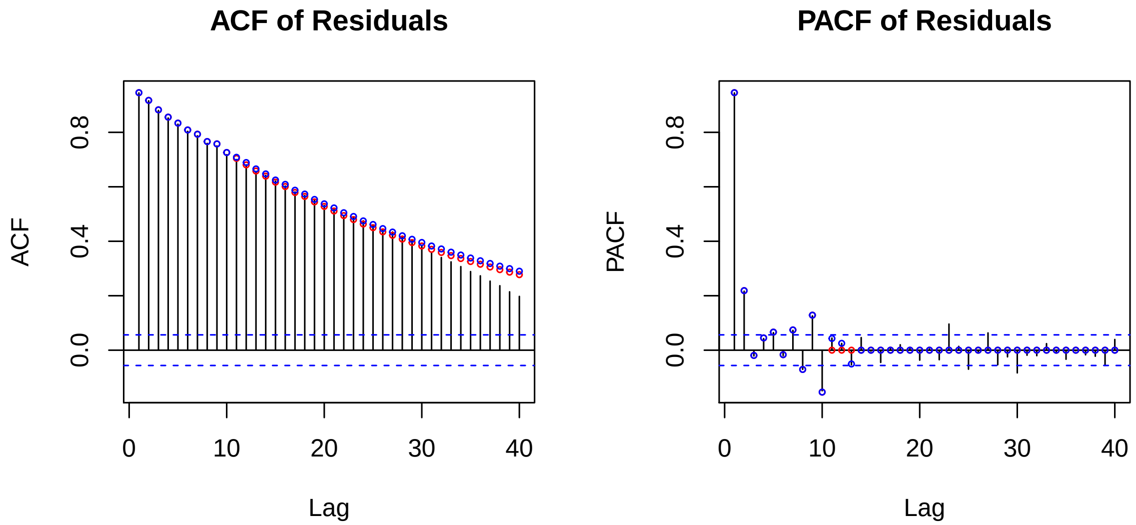

Figure 3: Autocorrelation function (ACF) and partial autocorrelation function (PACF) values for the parametric signal model's residuals. Red circles reflect the AR(10) model. Blue circles reflect the AR(13) model.

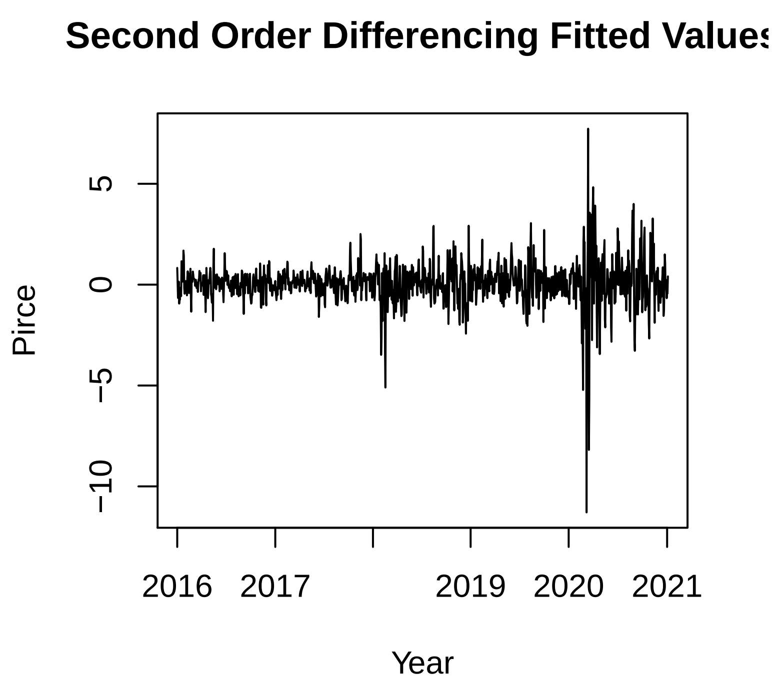

Figure 4: Diagnostics for differencing "signal model". The plot shows the differences themselves, to be assessed for stationarity.

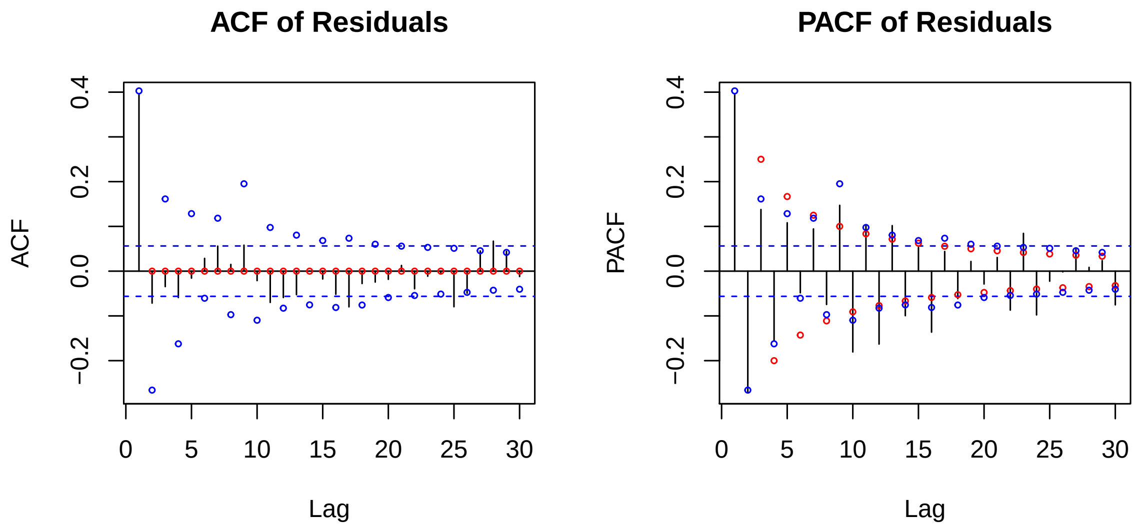

Figure 5: Autocorrelation function (ACF) and partial autocorrelation function (PACF) values for the differencing model. Red circles reflect the SMA(1)[7] model, while the blue circles reflect the MA(1).

3.2.2 Differencing with MA(10)

The sample ACF and PACF for the second order differences are shown in Figure @ref(fig:acf2). A minor seasonality of 10, or biweekly, is shown in the ACF plot. As a result, a MA(10) model can be applied here. The blue circles fitts well on the PACF plot as the sample PACF varies.

4 Model Compaison and Selection

Table 1: Cross-validated out-of-sample root mean squared prediction error for the four models under consideration.

	RMSPE
Parametric Model + AR(10)	1.799547
Parametric Model + AR(13)	1.815922
Differencing + MA(2)	1.849691
Differencing + MA(10)	1.842339

5 Results

To forecast the stock price, a parametric model of time will be used. Let Price_t be thestock price on day t with additive noise term X_t, which is restated as part of equation below. X_t is a stationary process defined by AR(10), where W_t is white noise with variance σ_W^2.

$$\text{Price}_t = \beta_0 + \beta_1 t + \beta_2 I_{\text{eighteen}_t} + \beta_3 t I_{\text{eighteen}_t} + \beta_4 I_{\text{covid}_t} + X_t$$

There are several binary indicators in this model. I_{eighteen_t} indicates if day t is in 2018 and early 2019. I_{covid_t} indicates if day t is under the impact of the stock market crash in March 2020. ϕ, Φ, θ, and all of the β's are coefficients that will be estimated in the next subsection.

5.1 Estimation of model parameters

Estimates of the model parameters are given in Table 2 in Appendix 1. Note that the stock price increases gradually, but the year of 2018 and the stock crash that had huge impact on the market also affected LOSI's price.

5.2 Prediction

Figure @ref(fig:forecasts) shows the forecasted values of sales for the first ten trading days of 2021. The model

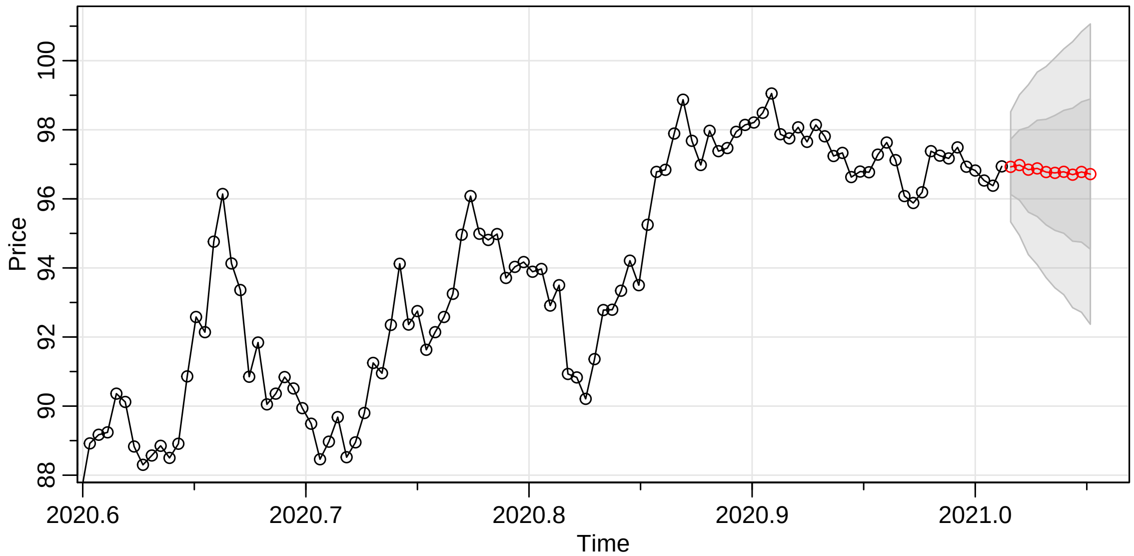

Figure 6: Forecasts of air conditioner sales for Chill-E-AC. The x-axis is time in years. The black points are the recent historical sales data. The red points are the forecasts for the first ten days in July 2019. The dark/light grey bands are the one/two standard error bands, representing 68%/95% prediction intervals, respectively. [Note to students: the x-axis would ideally be dates, but currently it is time in whole years, which is better than just integers "t". Sarima.for is proving to be a challenge to customize with their options; of course, I/you could just take their code and customize it correctly.]

6 Appendix 1 - Table of Parameter Estimates

Parameter	Estimate	SE	Coefficient Description [not required]
β_0	45.367	0.169	Intercept
β_1	0.038	0.000	Time
β_2	-2.549	0.978	Year 2018
β_3	-0.004	0.001	Time $\times$ year 2018 interaction
β_4	-7.347	0.632	Covid impact
ϕ	0.108	0.45	AR coefficient