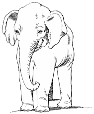
Long Memory ARMA: ARFIMA(p, d, q)

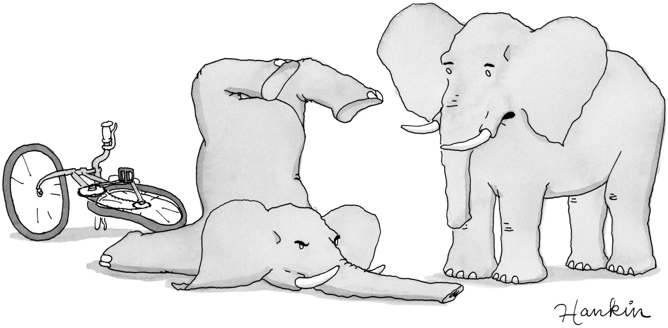
Sara Bakhtiari Kelsey Cherland
Jessica Huey Jessica Romero

California State University, Fullerton

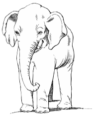
August 2021



An Elephant Never Forgets: Go Titans!



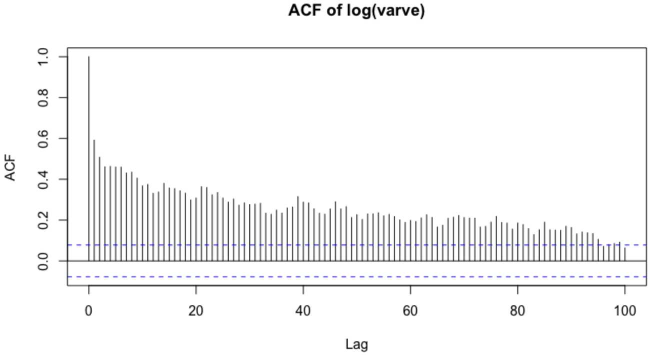
“Once you learn, though, you’ll never forget.”

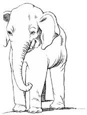


Warm-up Example: Glacial Varve

- Question for Audience: Is this ACF plot stationary?

  Enter "Yes"or "No"in the chat.

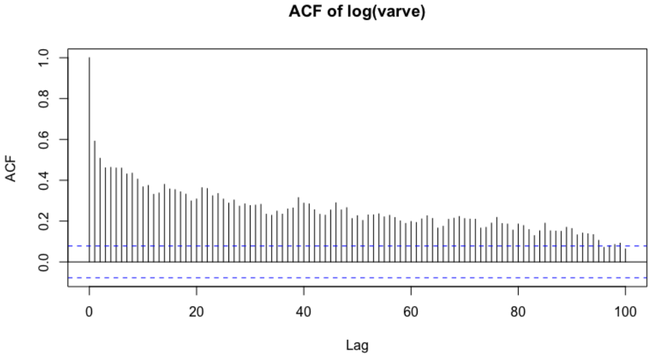


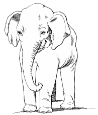


Warm-up Example: Glacial Varve

- Answer: This ACF plot is not stationary.

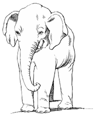
This means that past values are highly correlated with future values for almost 100 lags. When the decay of autocorrelation is so gradual, we may consider Long Memory ARMA or ARFIMA(p,d,q).





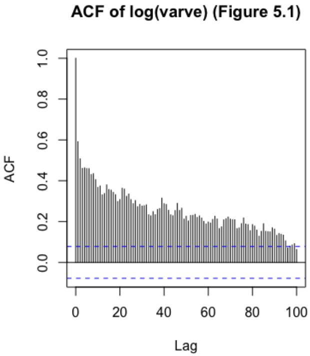
Introducing Long Memory ARMA

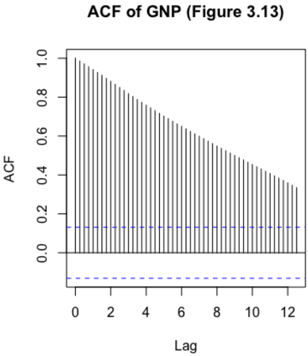
- Applications occur in hydrology, environmental data, and economic data
- Sample autocorrelations persist for a long time
- We have learned that to improve stationarity, we can use differencing:
$$(1 - B)^d x_t$$
- In long-memory, the ACF decays slowly despite being stationary and even d=1 might be *over-differencing*



When Are We in Long Memory?

- If the ACF decays "hyperbolically"and more slowly than a typical "short memory"ARMA process, we should consider long memory methodology
- The context is environmental data or economics/finance data





Introducing Long Memory ARMA

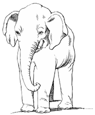
- Long memory time series methodology suggests that in order to deal with a persistent ACF, we should use fractional differencing
- In a short memory ARMA

$$\sum_{h=-\infty}^{\infty} |\rho(h)| < \infty$$

(finite)

- In long-memory, after fractional differencing:

$$\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty$$



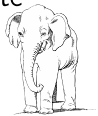
(infinite)

Introducing Long Memory ARMA

- Using the backward shift operations we are familiar with, the fractional differencing operator is defined by:

$$(1-B)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)B^j}{\Gamma(j+1)\Gamma(-d)}$$

  where d>-1 and $\Gamma(\cdot)$ is the standard gamma function (see eq. 5.2, 5.3)
- The ARFIMA model restricts d to $|d| < 0.5$ to avoid infinite variance.



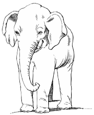
Introducing Long Memory ARMA

- The general ARFIMA(p, d, q) form, where -.5 < d < .5, can be written as

$$\phi(B)\nabla^d(x_t - \mu) = \theta(B)w_t$$

  where $\phi(B)$ and $\theta(B)$ are like what we have seen in Chap. 3.

- See equations (5.13)-(5.16) for a more complete understanding.

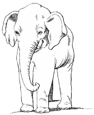


Introducing Long Memory ARMA

- For large h, the ACF of long memory time series can be expressed as:

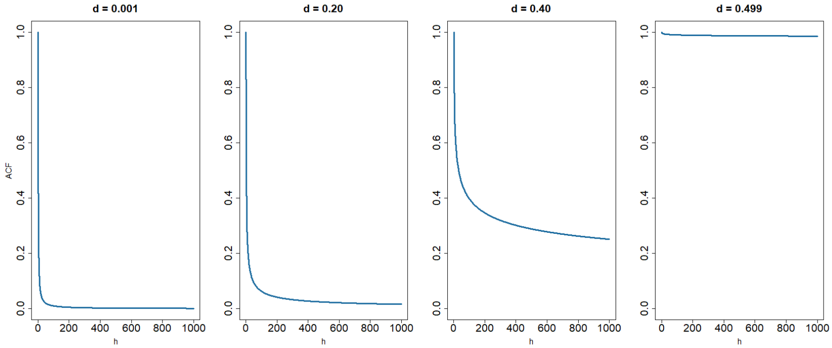
$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} \sim h^{2d-1}$$

- There are examples of this approximate distribution on the next slide.



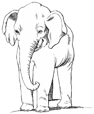
Introducing Long Memory ARMA

- As d approaches 0.5, the distribution of $\rho(h)$ becomes less stationary. We want d to be closer to 0.



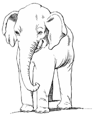
Introducing Long Memory ARMA

- For a long memory time series, the sum of $|\rho(h)|$ is equal to ∞ for 0<d<0.5.
- The value of d is unknown and must be estimated based on the data. Two common methods our textbook discusses for estimating the fractional difference, d, are the MLE estimation and the Newton-Raphson method of approximation.
- `fracdiff()` uses the MLE estimation method of approximation.



Introducing Long Memory ARMA

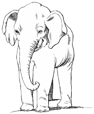
- Long memory methods of estimating d can yield different values of $\hat{d}$ and the standard errors
- The varve series example in our text shows that from `fracdiff()`, $\hat{d} = 0.384$ and s.e. $= 4.6 \times 10^{-6}$.
- Fitting the same model using frequency domain methods and the Whittle likelihood (4.85) provides an estimation of $\hat{d} = 0.380$ and s.e. $= .028$. The book highlights that the s.e. result is much more reasonable, but rather than dive further into a Whittle approximation, we will apply `fracdiff()` and demonstrate how we can apply the concept of fractional differencing.



Example: Dow Jones Industrial Average

Application Example

We will discuss an original analysis of the Dow Jones Industrial Average. This economic index was created and named after Charles Dow and Edward Jones, a business associate and statistician. The Dow Jones is a price-weighted average of 30 stocks. Its performance is widely considered a useful indicator of how the U.S. stock market is performing.

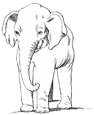


Example: Dow Jones Industrial Average-U.S.

Source

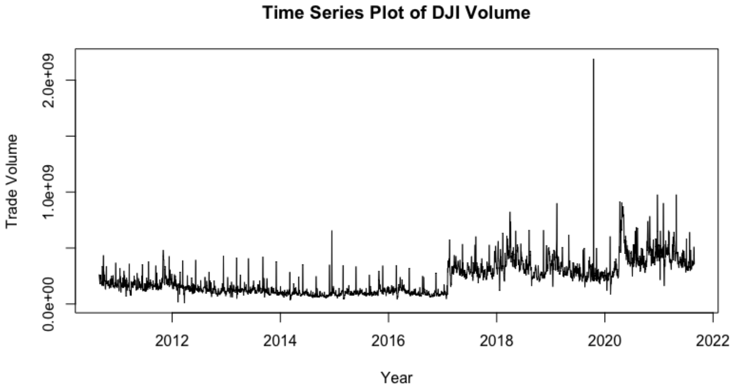
Available from: `https://stooq.com/q/d/?s=^dji`

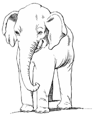
- Our data selection spans 135 months (5/27/2010 to 7/21/2021)
- We are measuring the ‘daily volume’ of shares traded within one day



Example: Dow Jones Industrial Average-U.S.

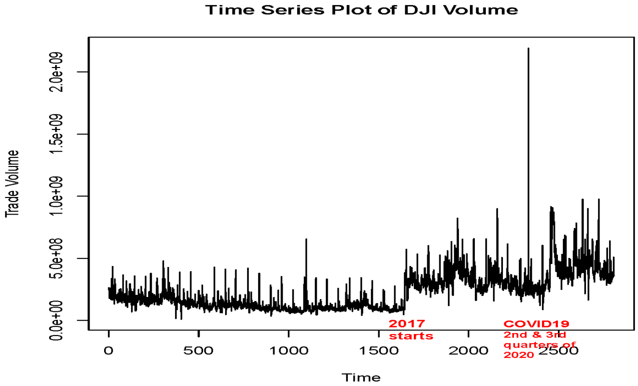
Question for Audience: Is the mean constant in this time series?

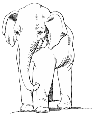




Example: Dow Jones Industrial Average

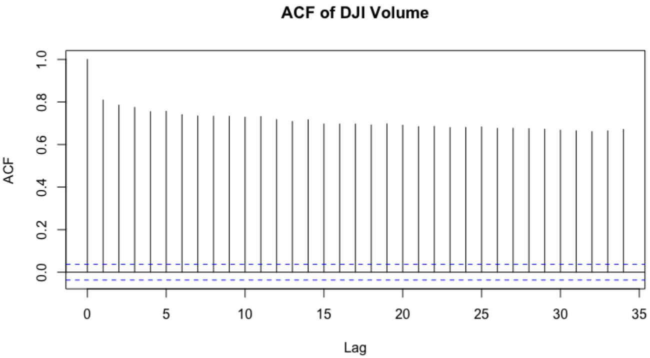
- Answer: The mean is not constant in this time series. The time series shows some erratic behavior between the days 2300 to 2470 (roughly)
- There are some BIG spikes after the start of year 2017

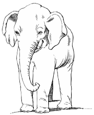




Example: Dow Jones Industrial Average

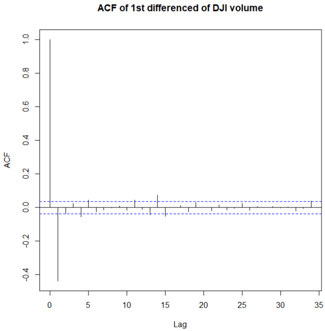
- The ACF dies down slowly, indicating non-stationarity and possible long memory phenomenon

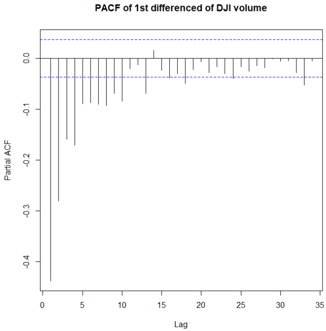


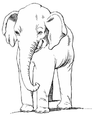


Example: Dow Jones Industrial Average

- Short memory methodology would say we should take a first difference
- Regular differencing (with d=1) improves stationarity
- An ARIMA(0,1,1) model seems reasonable (Table 3.1)

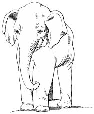




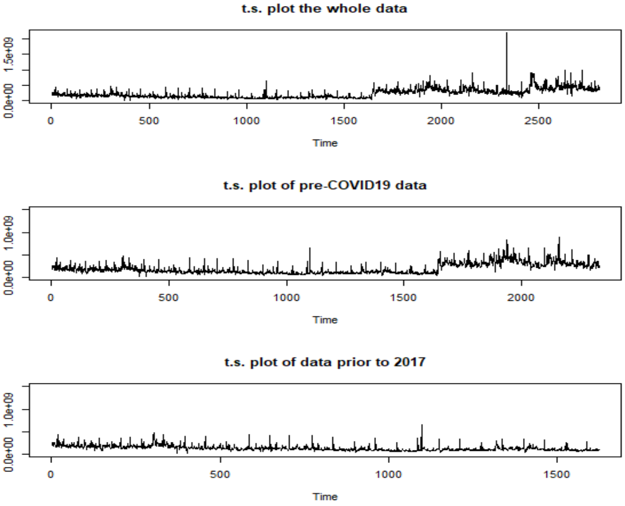


Example: Dow Jones Industrial Average

- Next, we will look at this from a long memory point of view
- We will estimate d for fractional differencing ($0 < |d| < 0.5$) for three time periods
- Three periods: (i) whole data (12 years), (ii) pre-COVID19 which runs to about December 2019, and (iii) data up to the end of 2016
- We estimate parameter d, using R package `fracdiff` for fractional differencing (along with the estimation of various model parameters).

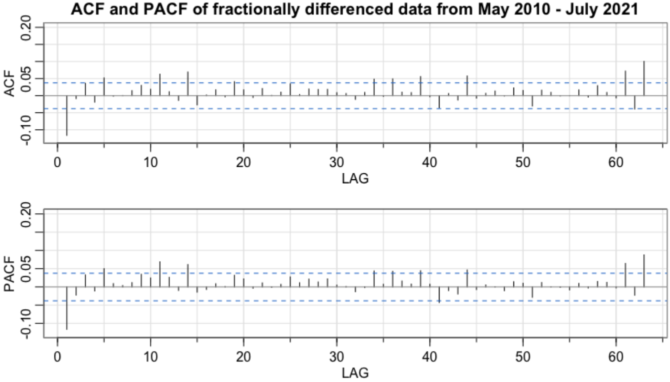


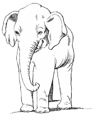
Dow Jones Example



(i) Dow Jones for May 2010 - July 2021

- These are the ACF and PACF plots for the whole dataset (May 2010 - July 2021) with best difference d.

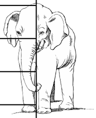




(i) Dow Jones for May 2010 - July 2021

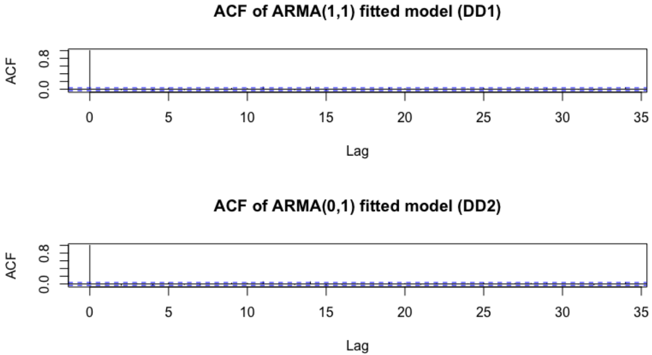
- Using `fracdiff`, we found d=0.4273039
- The ACF/PACF plots in conjunction with Table 3.1 suggests considering ARMA(1,1) and ARMA(0,1) as reasonable models for the fractionally differenced data
- The outputs show that ARMA(0,1) may be better (since the coefficient of AR is not significant).

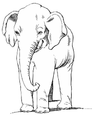
ARMA(1,1) DD1 summary				
	Estimate	Std. Error	z value	Pr(>|z|)
d	0.496989	0.004147	119.841	< 2e-16 ***
ar	0.109825	0.092040	1.193	0.232778 (not significant)
ma	0.297177	0.087982	3.378	0.000731 ***
ARMA(0,1) DD2 summary				
d	0.496377	0.004912	101.064	<2e-16 ***
ma	0.191033	0.019206	9.946	<2e-16 ***



(i) Dow Jones for May 2010 - July 2021

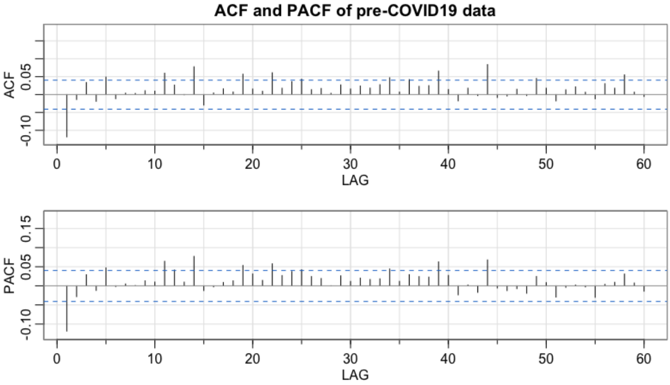
- The following residual plots of the fitted models ARMA(1,1) and ARMA(0,1) look like white noise.

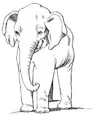




(ii) Dow Jones for Pre-Covid-19

- These are the ACF and PACF plots for the fractionally differenced pre-COVID19 data.



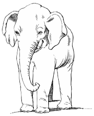


(ii) Dow Jones for Pre-Covid-19

- Using `fracdiff`, we found that $d = 0.3789942$
- The ACF/PACF plots in conjunction with Table 3.1 suggests considering ARMA(1,1) and ARMA(0,1) as reasonable models for the fractionally differenced data

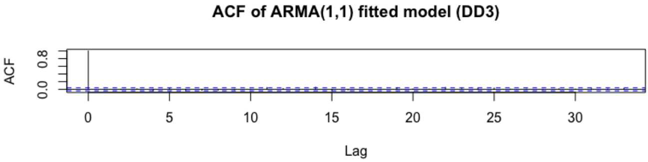
ARMA (1,1) DD3 summary				
	Estimate	Std. Error	z value	Pr(>|z|)
d	0.494450	0.007543	65.552	<2e-16 ***
ar	0.218041	0.078219	2.788	0.00531 **
ma	0.459774	0.071334	6.445	1.15e-10 ***

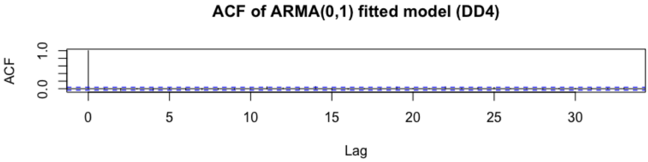
ARMA (0,1) DD4 summary				
	Estimate	Std. Error	z value	Pr(>|z|)
d	0.48654	0.01419	34.299	<2e-16 ***
ma	0.24519	0.02461	9.965	<2e-16 ***

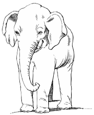


(ii) Dow Jones for Pre-Covid-19

- The residuals of the ARMA(1,1) and ARMA(0,1) models look like white noise:

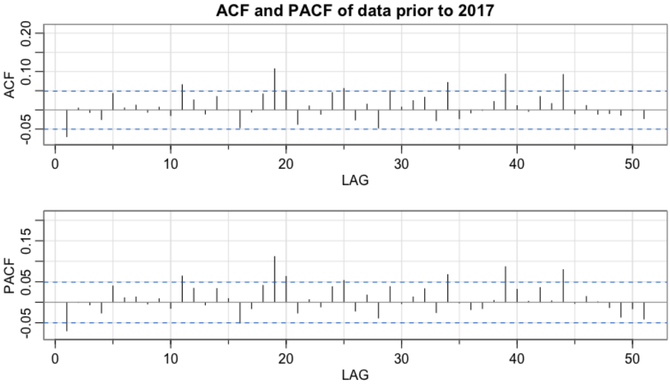


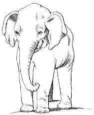




(iii) Dow Jones for Pre-2017

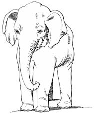
- These are the ACF and PACF plots for the fractionally differenced data up to the start of 2017.





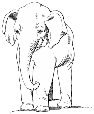
(iii) Dow Jones for Pre-2017

- Using `fracdiff`, we found that d=0.307969
- The ACF/PACF plots in conjunction with Table 3.1 suggest considering ARMA(1,1) and ARMA(0,1) as reasonable models for the fractionally differenced data
- We will compare all three time periods in our conclusion



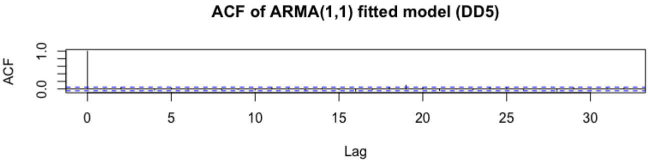
ARMA(1,1) DD5 summary				
	Estimate	Std. Error	z value	Pr(>|z|)
d	0.44550	0.02851	15.628	<2e-16 ***
ar	0.43478	0.07780	5.589	2.29e-08 ***
ma	0.64937	0.06452	10.065	< 2e-16 ***

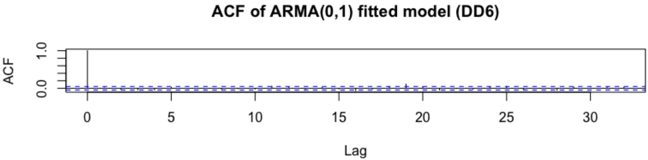
ARMA(0,1) DD6 summary				
	Estimate	Std. Error	z value	Pr(>|z|)
d	0.38257	0.02546	15.024	<2e-16 ***
ma	0.16025	0.03639	4.403	1.07e-05 ***

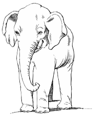


(iii) Dow Jones for Pre-2017

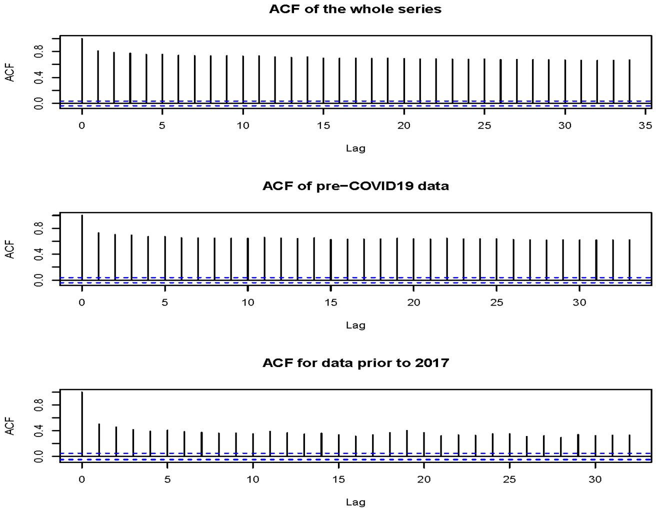
- The residuals of the ARMA(1,1) and ARMA(0,1) models look like white noise

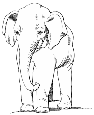






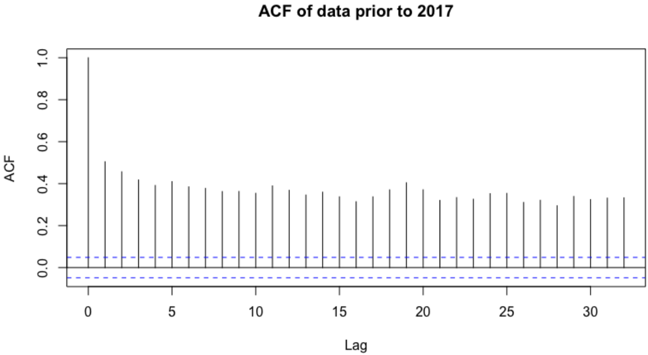
ACF plots for (i)-(iii)



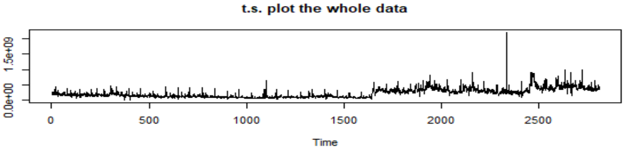


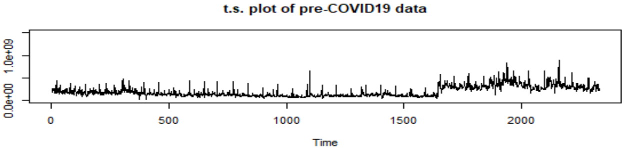
ACF plots for (i)-(iii)

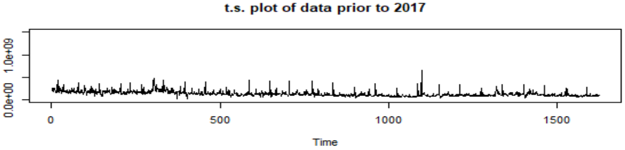
- Looking at the ACF plots of the three time series corresponding to (i), (ii), and (iii), one may consider (i) and (ii) as nonstationary series, whereas the data corresponding to (iii) is long-memory (but stationary-looking); see the time series plot of the whole data.

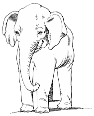


Conclusions









Conclusions

Remarks

Fractional differencing helps us fit models to long memory time series. The value of d, changes for different periods of volatility. The long memory analysis for parts (i) and (ii) are not the best choices because it is non-stationary. However, an ARFIMA(p, d, q) model is the perfect choice for part (iii).