# Autocovariance of ARMA

Jared Fisher

Lecture 6b

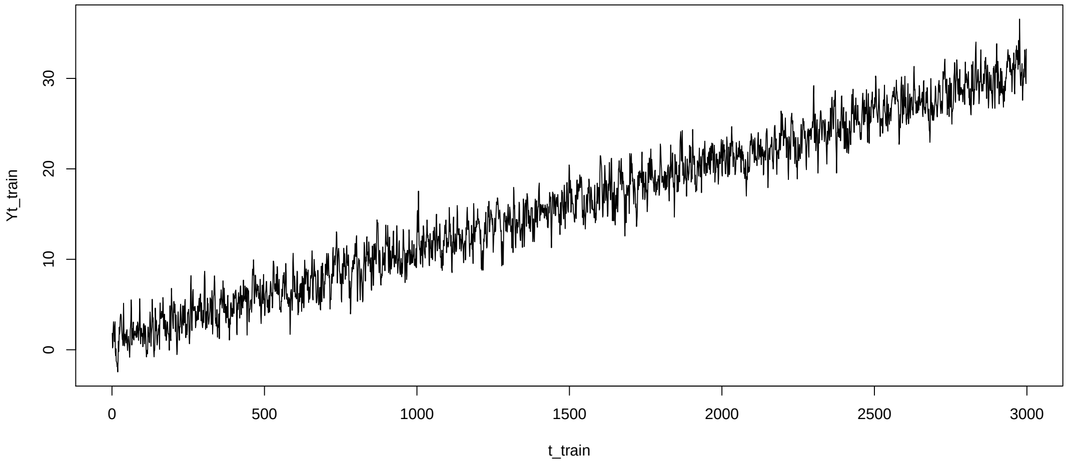
# Announcements

# Announcements

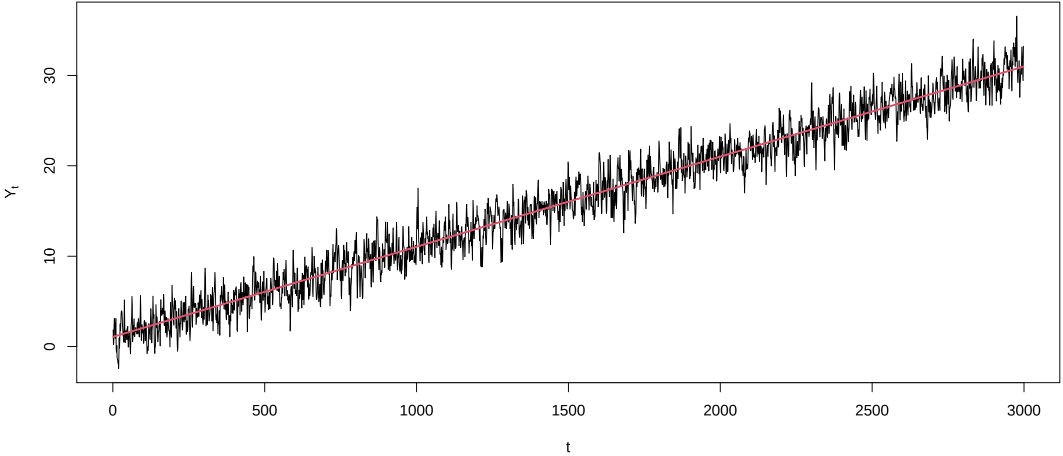
- Homework 4 is due Wednesday March 17 by 11:59pm
- Project Checkpoint 4 will be due Wednesday March 31 (the week after Spring Break)

# Recap

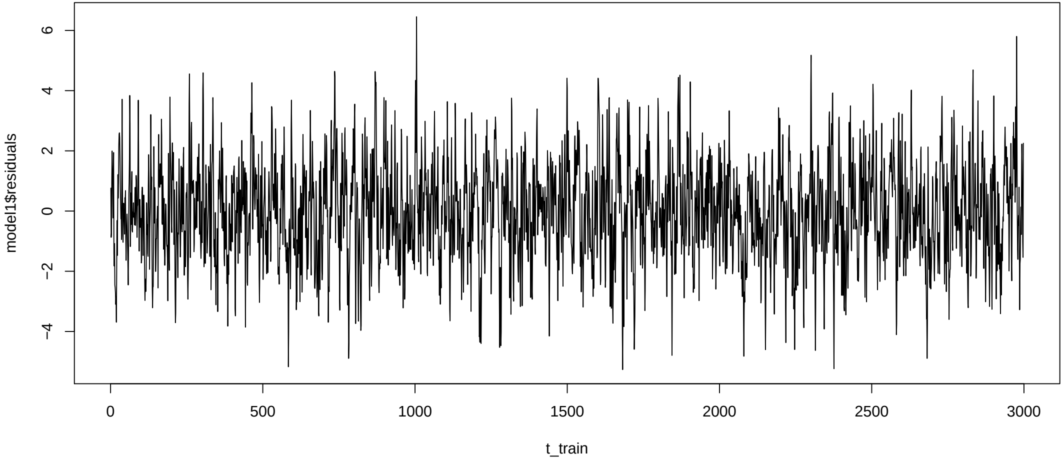
# Big Picture: modeling and forecasting



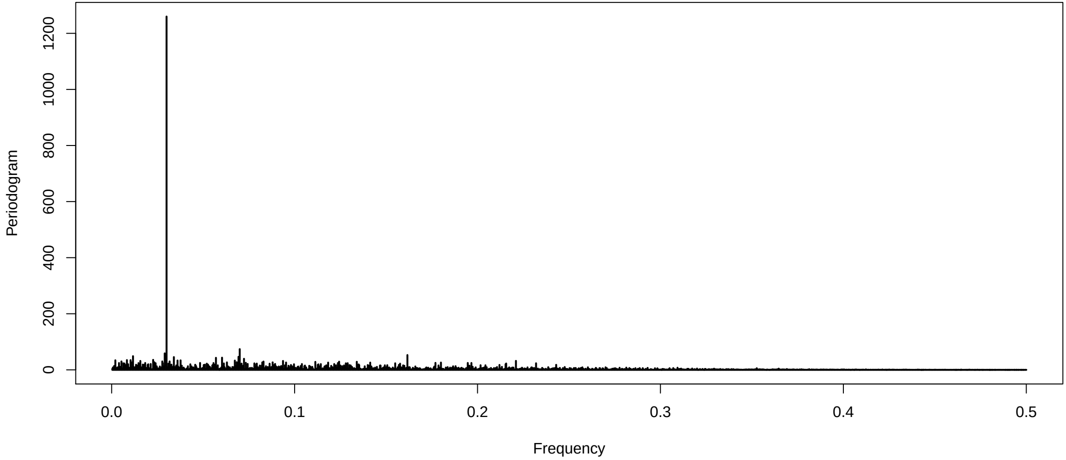
# Model the linear trend



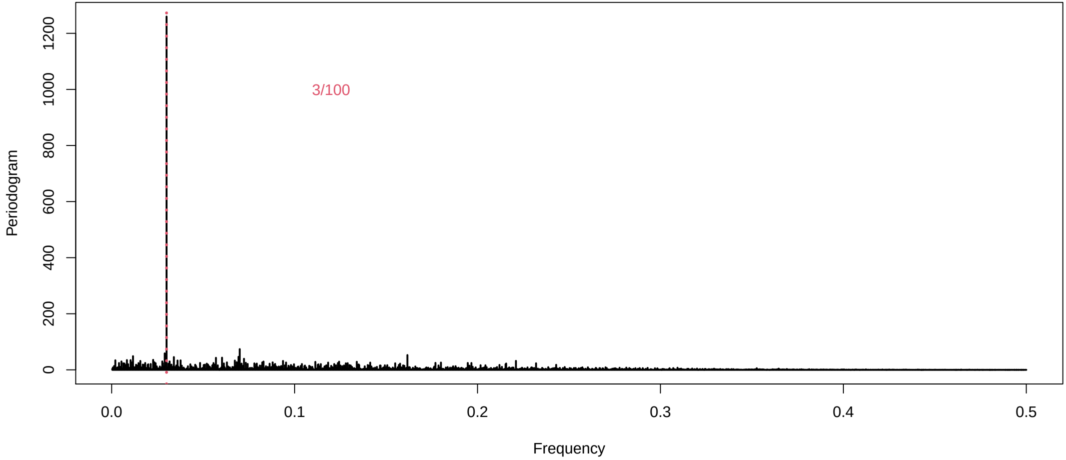
# Residuals with Trend removed



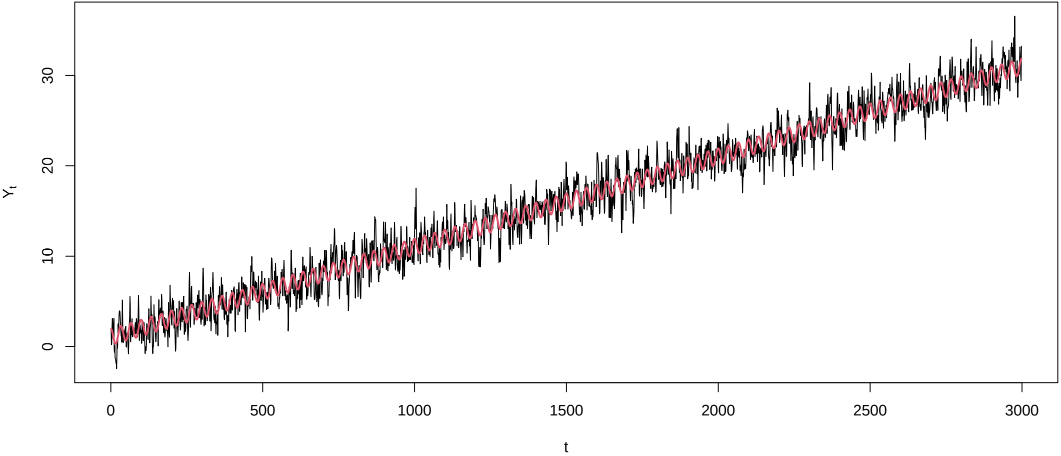
# No more trend, check periodogram for seasonality



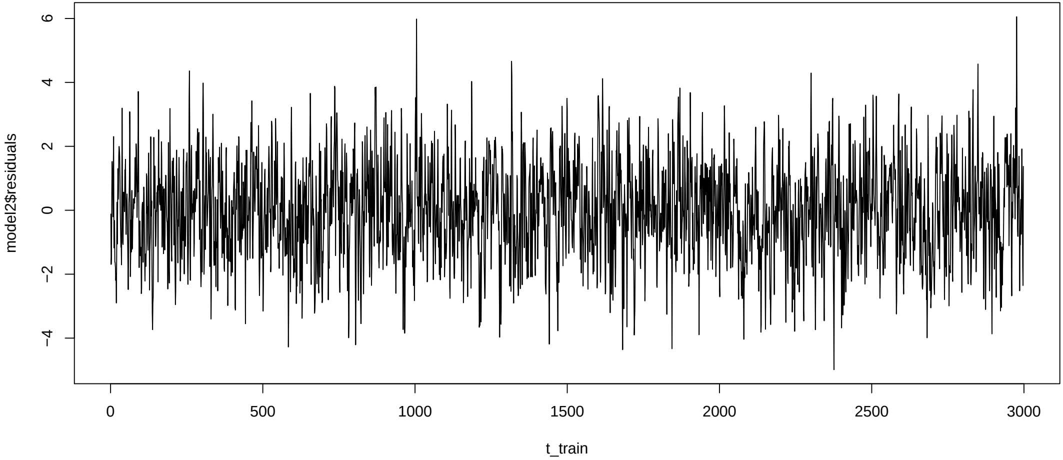
# Frequency is clearly 3/100



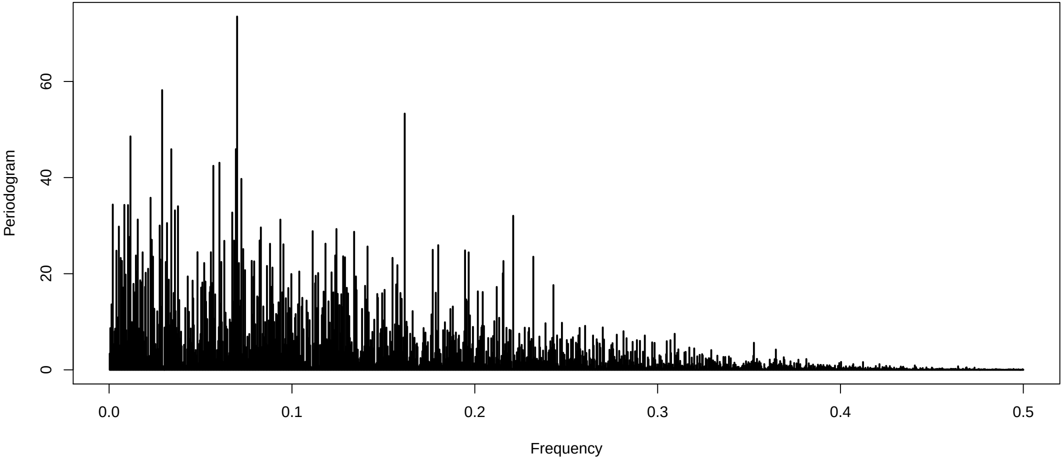
# Add Sinusoid to model



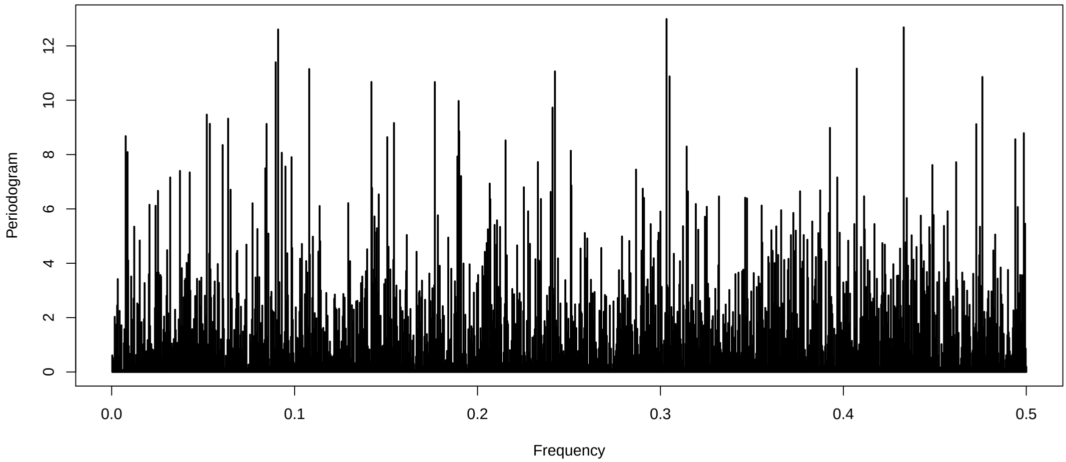
# Residuals without Linear Trend and Sinusoid



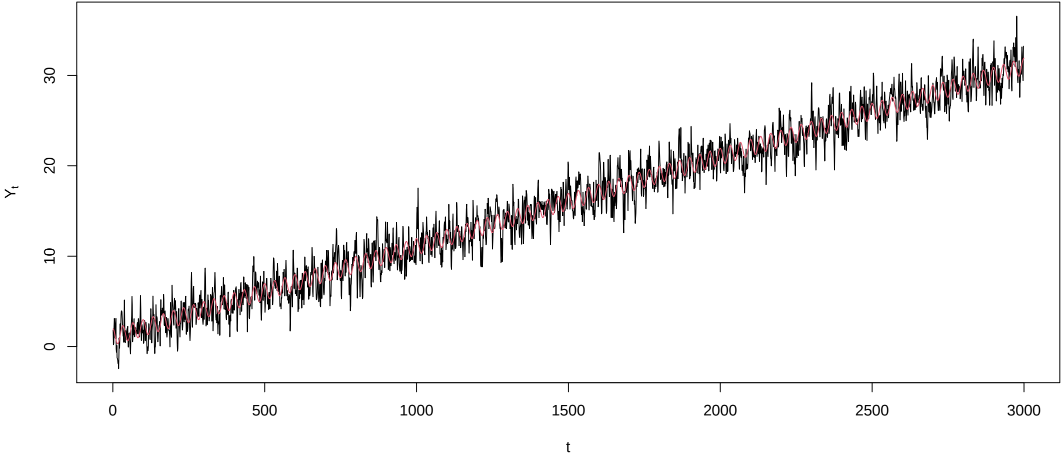
# No more large spikes either



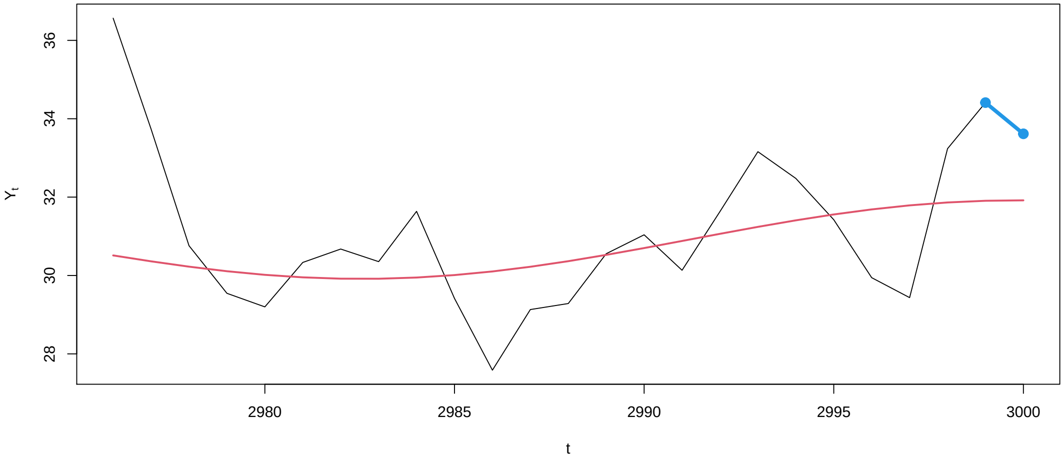
# For reference: Periodogram of Gaussian Noise



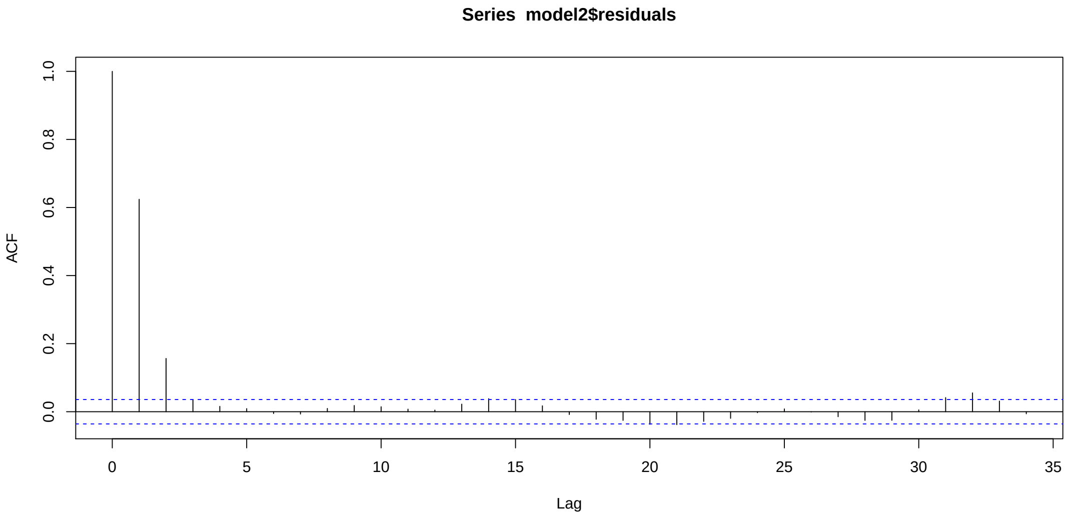
# Use this current model to forecast



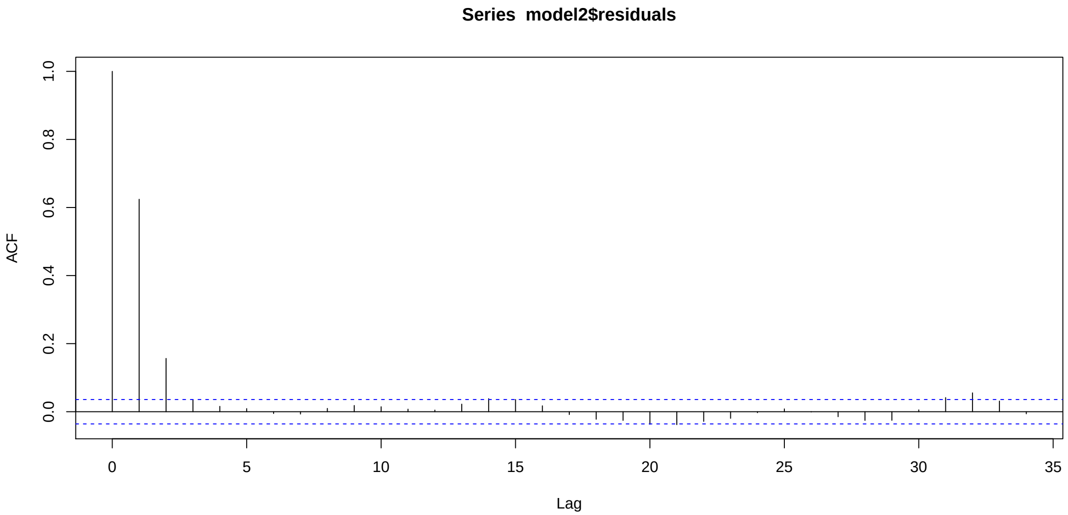
# Zoom in: Forecasting next two points, good not great?



# What's next, sinusoids? What does the ACF say?



# ACF plot begs the question: what does ARMA(p,q)'s ACF look like?



# Big Picture

- We're learning the ARMA(p,q) in order to use it for modeling a stationary process (e.g. residuals)
- In practice, after we have effectively pursued stationarity, we will need to pick p and q to appropriately model the remaining noise.
- After picking p and q, we'll need to estimate $\theta_1, ..., \theta_q$ and $\phi_1, ..., \phi_p$
- Right now we're building the machinery to do this, and know how it works!

# Today: Autocovariance function

- First, as we already assume ARMA(p,q) is zero-mean, we now need to understand the ACVF (autocovariance function), which is $\gamma_X(h)$.
- We can then construct the ACF (autocorrelation function) of ARMA(p,q) as $\rho_x(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$.
- Now a brief review of ARMA's important details

# ARMA(p,q)

Definition: A (zero mean) *autoregressive moving average* model of order $p$ and $q$ is of the form

$$\phi(B)X_t = \theta(B)W_t$$

where $\phi(B)$ is the AR operator, $\theta(B)$ is the MA operator, and $\{W_t\}$ is white noise.

# ARMA(p,q)

- Expanding the operators:

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + ...\theta_q W_{t-q}$$

- Rearranged for forecasting:

$$X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + ...\theta_q W_{t-q}$$

- Side note: R's functions expect ar=c($\phi_1$,$\phi_2$,. . . ) and ma=c($\theta_1$,$\theta_2$,. . . )

## Invertibility and Causality

- As MA and AR models are special cases of ARMA models, the invertible and causal conditions from MA and AR models, respectively, carry over the ARMA models.

# Invertibility

- An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$ is said to be **invertible** if $\theta(z) \neq 0$ for any $|z| \leq 1$
- Equivalently, an ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$ is **invertible** if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$W_t = \pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

as with inveribility theorem for MA(q), where $\pi(z)$ can be determined by solving

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1,$$

# Causality

- An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$ is said to be **causal** if $\phi(z) \neq 0$ for any $|z| \leq 1$.
- We can equivalently say ARMA(p,q) is **causal** if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$X_t = \psi(B)W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

as with the causality theorem for AR(p), where $\psi(z)$ can be determined by solving

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1,$$

# Unique Stationary Solution

An ARMA(p,q) model $\phi(B)X_t = \theta(B)W_t$ has a unique **stationary solution** if and only if $\phi(z) \neq 0$ for any $|z| = 1$

# Alignment with book

- For a few class periods, we've been in TSA4e's section 3.1
- Now we dive into section 3.2

# Where we are at now

- Note two major needs in the curriculum we've built so far:

1. Theory: we want to solve for the unique stationary solution of a given problem, which in the AR lecture we showed is in the form

$$X_t = \psi(B)W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$$

2. Application and theory: we need to be able to obtain the ACF of a given ARMA(p,q). Recall that we know a convenient form of the ACVF for the $MA(\infty)$ model

$$\gamma_X(h) = \sigma_W^2 \sum_{j=0}^{\infty} \theta_j \theta_{j+h}$$

- We can solve both problems with the same approach: obtain the unique stationary solution of an ARMA(p,q) and its autocovariance/autocorrelation.

# Step through the idea:

- Given p and q, we have $\phi(B)X_t = \theta(B)W_t$, but we want $X_t = \psi(B)W_t$, such that $\psi(B) = \frac{\theta(B)}{\phi(B)}$.
- First, solve for the values of $\psi_j$.
- Second, calculate $\gamma_X(h)$, the autocovariance function of $X_t$. (With this we can easily calculate the ACF $\rho_X(h)$).
- Now for a brief demonstration of the ideas on the "whiteboard"

# Whiteboard

# Solutions

# Finding $\psi$ = Finding the *Unique* Causal Solution

- Recall the "difference equation" of interest is $\phi(B)X_t = \theta(B)W_t$
- Our desired solution is in the form $X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$
- We know and understand $\{W_t\}$, thus to find the solution, we need to find $\psi_j$, $j > 0$ (as $\psi_0 = 1$)
- In order to find the unique causal solution $X_t = \psi(B)W_t$ (if it exists), we need to divide the two polynomials

$$\frac{\theta(z)}{\phi(z)} = \psi(z)$$

.

# Dividing Polynomials

In order to find the unique causal solution $X_t = \psi(B)W_t$ (if it exists), we need to divide the two polynomials $\theta(z)/\phi(z) = \psi(z)$. My main approach:

Because $\psi(z) = \theta(z)/\phi(z)$, we have

$$(1 - \phi_1 z - \cdots - \phi_p z^p)(\psi_0 + \psi_1 z + \dots) = 1 + \theta_1 z + \cdots + \theta_q z^q.$$

Equate the coefficients of $z^j$ on both sides for $j = 0, 1, 2, \dots$ to get

$$\begin{aligned}
\psi_0 &= 1 \\
\psi_1 &= \theta_1 + \psi_0\phi_1 \\
\psi_2 &= \theta_2 + \psi_1\phi_1 + \psi_0\phi_2 \\
\psi_3 &= \theta_3 + \phi_1\psi_2 + \phi_2\psi_1 + \phi_3\psi_0 \\
&\vdots
\end{aligned}$$

# Example: $X_t - 0.5X_{t-1} = W_t + 0.4W_{t-1}$

- Consider the following ARMA(1, 1) difference equation above, where $\{W_t\}$ is white noise. Is it invertible? Is it causal? Does this have a unique stationary solution? Find the solution.
- The moving average polynomial is $\theta(z) = 1 + 0.4z$.
- This has the root $2.5 > 1$, so it is invertible.

## Example: $X_t - 0.5X_{t-1} = W_t + 0.4W_{t-1}$

- The autoregressive polynomial is $\phi(z) = 1 - 0.5z$.
- $\phi$ has only one root: $z = 2$.
- Root $z \neq 1 \implies \exists$ a unique stationary solution.
- Root $|z| > 1 \implies$ the unique stationary solution is causal.

## Example: $X_t - 0.5X_{t-1} = W_t + 0.4W_{t-1}$

- Note $\phi_1 = 0.5$ and $\theta_1 = 0.4$
- Thus:

$$
\begin{aligned}
\psi_0 &= 1 \\
\psi_1 &= \theta_1 + \psi_0\phi_1 = 0.4 + (1)(0.5) = 0.9 \\
\psi_2 &= \theta_2 + \psi_1\phi_1 + \psi_0\phi_2 = 0 + 0.9 * 0.5 + 0 = 0.45 \\
\psi_3 &= \theta_3 + \psi_2\phi_1 + \psi_1\phi_2 + \psi_0\phi_3 = 0 + 0.45 * 0.5 + 0 + 0 = .225 \\
&\vdots \\
\psi_k &= \psi_{k-1}\phi_1 = 0.9(0.5^{k-1}), \text{ for } k > 0 \\
&\vdots
\end{aligned}
$$

- Remember this $\psi_k$, it will come up again in a minute.

# An alternate approach for dividing polynomials

(I don't use this method, but it could be useful in some cases)

Another way is to write $\phi(z) = (1 - a_1 z)(1 - a_2 z)\dots(1 - a_p z)$ where $1/a_1, \dots, 1/a_p$ are the (possibly complex) roots of $\phi(z)$ each satisfying $|a_i| < 1$ so that

$$\begin{aligned}
\psi(z) &= \frac{\theta(z)}{\phi(z)} \\
&= \frac{\theta(z)}{(1 - a_1 z)\dots(1 - a_p z)} \\
&= \theta(z)(1 - a_1 z)^{-1}\dots(1 - a_p z)^{-1} \\
&= \theta(z)(1 + a_1 z + a_1^2 z^2 + \dots)\dots(1 + a_p z + a_p^2 z^2 + \dots).
\end{aligned}$$

The product above can be multiplied out.

# Autocovariance Function of ARMA(p,q)

# Autocovariance Function (ACVF)

- Recall that causal ARMA series have the nice $MA(\infty)$ representation:

$$X_t = \psi(B)W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

- Also recall that $MA(\infty)$ has a convenient ACVF:

$$\gamma_X(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$$

# Autocovariance of ARMA Processes

- From now on, we will consider causal, stationary and invertible ARMA processes
- To get ARMA's ACF (autocorrelation function), we first need to find $\gamma_X(h)$, the ACVF (as the ACF can be calculated via $\rho_X(h) = \gamma_X(h)/\gamma_X(0)$).
- $\phi(B)X_t = \theta(B)W_t$ with $\phi(z) \neq 0$ for any $|z| \leq 1$ and $\theta(z) \neq 0$ for any $|z| \leq 1$.
- Here we'll discuss two different approaches on how to determine the ACVF $\gamma(h)$ and the ACF $\rho(h)$ for such a process.

# 1st Approach: Dividing Polynomials

# Dividing Polynomials: ARMA(1,1) Example

- This method requires calculation of the function $\psi(z)$ by dividing the two polynomials $\theta(z)$ and $\phi(z)$. Once we have the $\psi_j$ values, we plug them into the ACVF (which has $MA(\infty)$ form).
- Return to the example $X_t - 0.5X_{t-1} = W_t + 0.4W_{t-1}$ and plug in the values of $\psi_j$ that we already found

$$
\begin{aligned}
\gamma_X(h) &= \sigma_W^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h} \\
&= \sigma_W^2 \left( \psi_0 \psi_h + \sum_{j=1}^{\infty} \psi_j \psi_{j+h} \right) \\
&= \sigma_W^2 \left( \psi_h + \sum_{j=1}^{\infty} (0.9 * 0.5^{j-1})(0.9 * 0.5^{j+h-1}) \right) \\
&= \sigma_W^2 \left( \psi_h + 0.81 \sum_{j=1}^{\infty} 0.5^{2j+h-2} \right)
\end{aligned}
$$

# Dividing Polynomials: ARMA(1,1) Example

$$
\begin{aligned}
\gamma_X(h) &= \sigma_W^2 \left( \psi_h + 0.81 \sum_{j=1}^{\infty} 0.5^{2j+h-2} \right) \\
&= \sigma_W^2 \left( \psi_h + (0.81 * 0.5^{h-2}) \sum_{j=1}^{\infty} 0.25^{j} \right) \\
&= \sigma_W^2 \left( \psi_h + 0.81 * 0.5^{h-2} * \frac{1/4}{3/4} \right) \\
&= \sigma_W^2 \left( \psi_h + 0.27 * 0.5^{h-2} \right)
\end{aligned}
$$

# Dividing Polynomials: ARMA(1,1) Example

- So

$$\begin{aligned}\gamma_X(h=0) &= \sigma_W^2 \left(\psi_0 + 0.27 * 0.5^{0-2}\right) \\ &= \sigma_W^2 (1 + 0.27 * 4) \\ &= \sigma_W^2 (2.08)\end{aligned}$$

- And for $h > 0$

$$\begin{aligned}\gamma_X(h) &= \sigma_W^2 \left(0.9 * 0.5^{h-1} + 0.27 * 0.5^{h-2}\right) \\ &= \sigma_W^2 * 0.5^h (2.88)\end{aligned}$$

- Which yields, for $h > 0$,

$$\begin{aligned}\rho_X(h) &= \gamma_X(h)/\gamma_X(0) \\ &= (\sigma_W^2 * 0.5^h (2.88))/(\sigma_W^2 (2.08)) \\ &= \frac{2.88}{2.08} 0.5^h\end{aligned}$$

# Dividing Polynomials: MA(2)

- For any MA process we have that $\psi(z) = \theta(z)$.
- For MA(2): $X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2}$, we have

$$\psi_0 = 1, \psi_1 = \theta_1, \psi_2 = \theta_2$$

  and $\psi_j = 0$ for $j \geq 3$.
- $\gamma_X(h) = \sigma_W^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$ yields:

$$\begin{aligned}
\gamma_X(0) &= \sigma_W^2(1 + \theta_1^2 + \theta_2^2) \\
\gamma_X(1) &= \sigma_W^2 \theta_1 (1 + \theta_2) \\
\gamma_X(2) &= \sigma_W^2 \theta_2 \\
\gamma_X(h) &= 0 \text{ for } h \geq 3.
\end{aligned}$$

# Dividing Polynomials: MA(2)

- The corresponding autocorrelations are given by

$$\rho_X(1) = \frac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2}$$
$$\rho_X(2) = \frac{\theta_2}{1+\theta_1^2+\theta_2^2}$$
$$\rho_X(h) = 0 \quad \text{for } h \geq 3.$$

# Dividing Polynomials: AR(1)

- AR(1): $X_t - \phi X_{t-1} = W_t$.
- If stationary and causal ( $|\phi| < 1$) $X_t = W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \ldots$.
- For our purposes, $\psi_j = \phi^j$ for $j = 0, 1, 2, \ldots$.
- Thus the ACVF:

$$\gamma_X(h) = \sigma_W^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+h} = \sigma_W^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} = \sigma_W^2 \frac{\phi^h}{1 - \phi^2} \text{ for } h > 0$$

- And the ACF: $\rho_X(h) = \phi^h$ for $h > 0$.

# 2nd Approach: Solve the (Covariance of) Difference Equations

# Solve the (Covariance of) Difference Equations

- For any ARMA(p,q) process $\{X_t\}$ with p>1 the difference equation is $\phi(B)X_t = \theta(B)W_t$. (If p=0, then it's MA(q) we've already found it's ACVF).
- Let $k \geq 0$; we have that

$$cov(\phi(B)X_t, X_{t-k}) = cov(\theta(B)W_t, X_{t-k}).$$

- For the left hand side

$$\begin{aligned}
&cov(\phi(B)X_t, X_{t-k}) \\
&= cov(X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p}, X_{t-k}) \\
&= cov(X_t, X_{t-k}) - \phi_1 cov(X_{t-1}, X_{t-k}) - \ldots - \phi_p cov(X_{t-p}, X_{t-k}) \\
&= \gamma_X(k) - \phi_1 \gamma_X(k-1) - \cdots - \phi_p \gamma_X(k-p).
\end{aligned}$$

## Solve Difference Equations

- For the right hand side, recall $X_t = \psi_0 W_t + \psi_1 W_{t-1} + \dots$ to get

$$
\begin{aligned}
&\operatorname{cov}(\theta(B) W_t, X_{t-k}) \\
&= \operatorname{cov}(W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}, \psi_0 W_{t-k} + \psi_1 W_{t-k-1} + \dots) \\
&= \begin{cases} (\psi_0 \theta_k + \psi_1 \theta_{k+1} + \cdots + \psi_{q-k} \theta_q)\, \sigma_W^2 & \text{if } k \leq q \\ 0 & \text{if } k > q \end{cases}
\end{aligned}
$$

- Equating both sides, we get for all $k \geq 0$:

$$
\gamma_X(k) - \phi_1 \gamma_X(k-1) - \cdots - \phi_p \gamma_X(k-p) = c_k
$$

with

$$
c_k = \begin{cases} (\psi_0 \theta_k + \psi_1 \theta_{k+1} + \cdots + \psi_{q-k} \theta_q)\, \sigma_W^2 & \text{for } 0 \leq k \leq q \\ 0 & \text{for } k > q. \end{cases}
$$

- Solve to get the autocovariance function $\gamma_X$

# But how exactly do we solve this equation?

- Note that $\gamma(k) = \gamma(-k)$
- We can build a linear system of $p + 1$ equations which we can solve for the $p + 1$ unknowns $\gamma(0), \ldots, \gamma(p)$.
- Then for $k > p$ we can recursively compute

$$\gamma_X(k) = c_k + \phi_1 \gamma_X(k-1) + \cdots + \phi_p \gamma_X(k-p).$$

# Example: Yule-Walker Equations

- Let q=0, i.e. look at AR(p) processes
- $c_k$ in has a very simple form (recall that $\psi_0 = \theta_0 = 1$), namely

$$c_k = \begin{cases} \sigma_W^2 & \text{if } k = 0 \\ 0 & \text{if } k > 0. \end{cases}$$

- These will also be useful later in the course, when we discuss estimation of the AR coefficients $\phi_1, \ldots, \phi_p$ from data.

# Example: Yule-Walker Equations with $p = 2$

- For the system of equations, we have

$$\gamma_X(k) - \phi_1\gamma_X(k-1) - \phi_2\gamma_X(k-2) = c_k$$

  and

$$c_k = \begin{cases} \sigma_W^2 & \text{if } k = 0 \\ 0 & \text{if } k > 0. \end{cases}$$

- System (with k=0,1,2):

$$\begin{bmatrix} \gamma_X(0) - \phi_1\gamma_X(0-1) - \phi_2\gamma_X(0-2) \\ \gamma_X(1) - \phi_1\gamma_X(1-1) - \phi_2\gamma_X(1-2) \\ \gamma_X(2) - \phi_1\gamma_X(2-1) - \phi_2\gamma_X(2-2) \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$

# Example: Yule-Walker Equations with $p = 2$

- Recall the ACVF property $\gamma(k) = \gamma(-k)$:

$$\begin{bmatrix} \gamma_X(0) - \phi_1\gamma_X(1) - \phi_2\gamma_X(2) \\ \gamma_X(1) - \phi_1\gamma_X(0) - \phi_2\gamma_X(1) \\ \gamma_X(2) - \phi_1\gamma_X(1) - \phi_2\gamma_X(0) \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$

- Insert definition of $c_k$:

$$\begin{bmatrix} \gamma_X(0) - \phi_1\gamma_X(1) - \phi_2\gamma_X(2) \\ \gamma_X(1) - \phi_1\gamma_X(0) - \phi_2\gamma_X(1) \\ \gamma_X(2) - \phi_1\gamma_X(1) - \phi_2\gamma_X(0) \end{bmatrix} = \begin{bmatrix} \sigma_W^2 \\ 0 \\ 0 \end{bmatrix}$$

# Example: Yule-Walker Equations with $p = 2$

- Seperate Unknowns:

$$\begin{bmatrix} 1 & -\phi_1 & -\phi_2 \\ -\phi_1 & 1-\phi_2 & 0 \\ -\phi_2 & -\phi_1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_X(0) \\ \gamma_X(1) \\ \gamma_X(2) \end{bmatrix} = \begin{bmatrix} \sigma_W^2 \\ 0 \\ 0 \end{bmatrix}$$

- Solve system:

$$\gamma_X(0) = \sigma_W^2 \frac{1-\phi_2}{(1+\phi_2)[(1-\phi_2)^2 - \phi_1^2]}$$

$$\gamma_X(1) = \sigma_W^2 \frac{\phi_1}{(1+\phi_2)[(1-\phi_2)^2 - \phi_1^2]}$$

$$\gamma_X(2) = \sigma_W^2 \frac{\phi_1^2 + \phi_2(1-\phi_2)}{(1+\phi_2)[(1-\phi_2)^2 - \phi_1^2]}$$

## Example: Yule-Walker Equations with $p = 2$

- For $k > 2$, recall

$$\gamma_X(k) = c_k + \phi_1 \gamma_X(k-1) + \cdots + \phi_p \gamma_X(k-p)$$

- Then for AR(2):

$$\gamma_X(k) = (0) + \phi_1 \gamma_X(k-1) + \phi_2 \gamma_X(k-2)$$

- k=3,4,...

$$\begin{aligned}
\gamma_X(3) &= \phi_1 \gamma_X(2) + \phi_2 \gamma_X(1) = \sigma_W^2 \frac{\phi_1^3 + \phi_1 \phi_2 (2 - \phi_2)}{(1+\phi_2)[(1-\phi_2)^2 - \phi_1^2]} \\
\gamma_X(4) &= \phi_1 \gamma_X(3) + \phi_2 \gamma_X(2) = \sigma_W^2 \frac{\phi_1^4 + \phi_1^2 \phi_2 (3 - \phi_2) + \phi_2^2 (1 - \phi_2)}{(1+\phi_2)[(1-\phi_2)^2 - \phi_1^2]} \\
&\vdots
\end{aligned}$$

# Example: Yule-Walker Equations with $p = 2$

Note that the autocorrelations look much simpler:

$$\begin{aligned}
\rho_X(1) =& \frac{\gamma_X(1)}{\gamma_X(0)} = \frac{\phi_1}{1-\phi_2} \\
\rho_X(2) =& \frac{\gamma_X(2)}{\gamma_X(0)} = \frac{\phi_1^2 + \phi_2(1-\phi_2)}{1-\phi_2} \\
&\vdots
\end{aligned}$$

# Notes on the Lab tomorrow:

- We'll try this for ARMA(1,1)
- Will we always get an explicit expression for $\gamma_X(k)$ when $k > 0$?
- Yes!
- One can obtain a closed form expression for $\gamma_X(k)$ which depends on the (possibly complex) zeros of the polynomial $\phi(z)$.

# Recap today

# Autocovariance of ARMA

- One way: solve for $\psi$ by dividing polynomials

$$(1 - \phi_1 z - \cdots - \phi_p z^p)(\psi_0 + \psi_1 z + \dots) = 1 + \theta_1 z + \cdots + \theta_q z^q.$$

- Another way: solve via the covariance of the difference equations

$$\gamma_X(k) - \phi_1 \gamma_X(k-1) - \cdots - \phi_p \gamma_X(k-p) = c_k$$

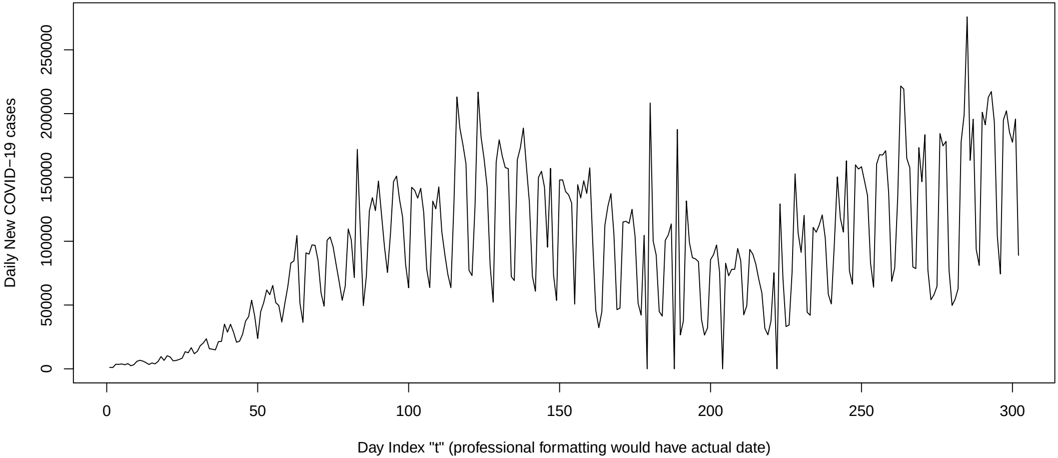
  with

$$c_k = \begin{cases} (\psi_0 \theta_k + \psi_1 \theta_{k+1} + \cdots + \psi_{q-k} \theta_q)\, \sigma_W^2 & \text{for } 0 \le k \le q \\ 0 & \text{for } k > q. \end{cases}$$

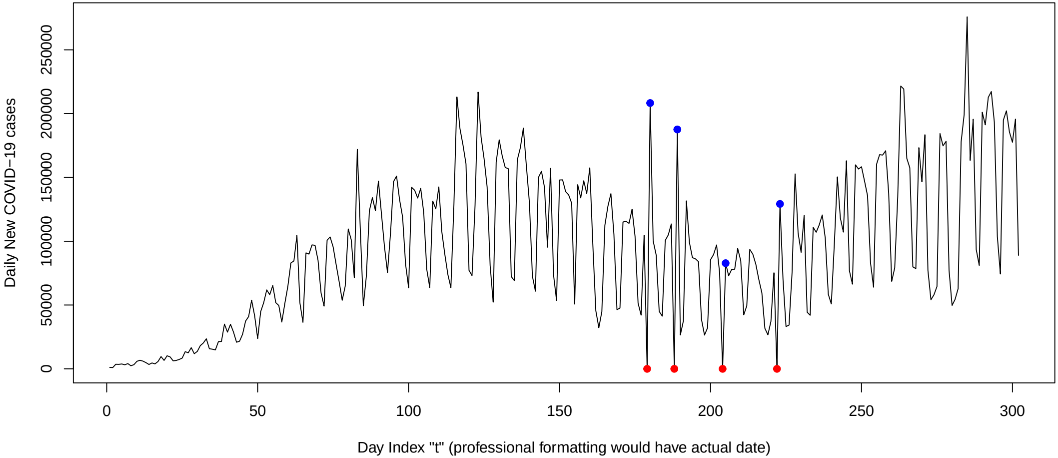
- See Example 3.7 and 3.8 in TSA4e for further examples.

# Thoughts on the Project

# Data issues? ("Real-world" data is messy ;)



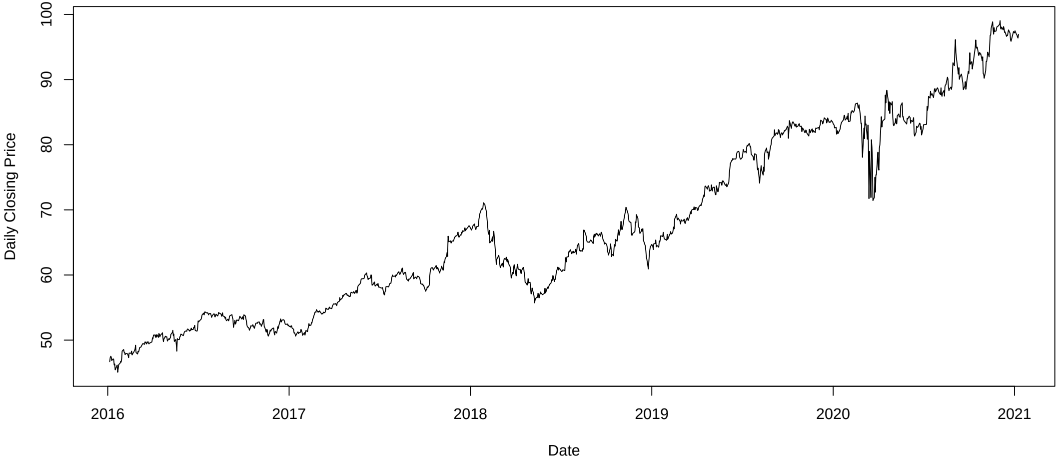
# Data issues? (“Real-world” data is messy ;)

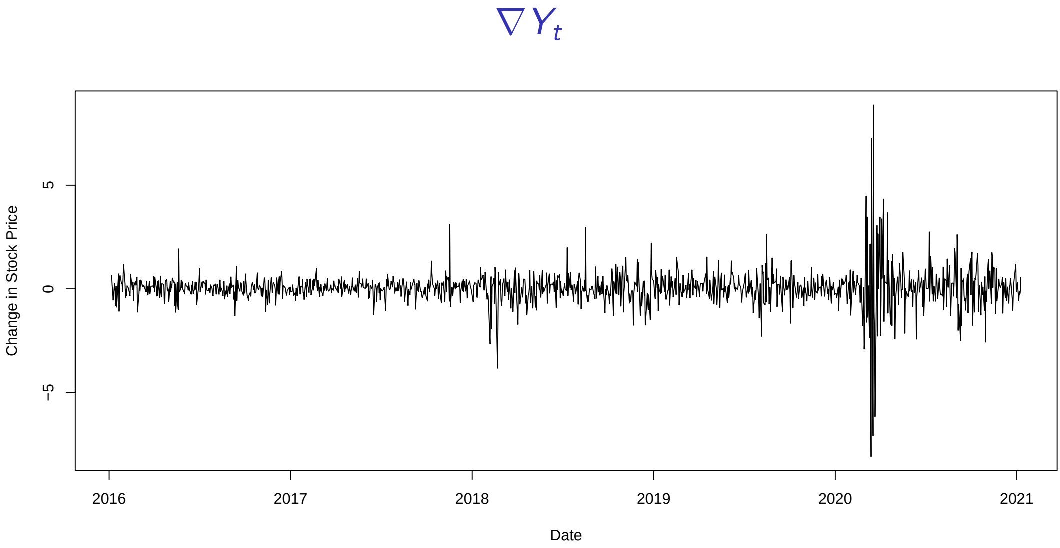


# Tips

- Remember that the end goal is accurate forecasts for $Y_{n+1}, ..., Y_{n+10}$. Our model should describe what we believe these points are going to do.
- Personally, I would not
  - Drop data points in the middle of my dataset. Why?
    - We might do this in typical stat/ML problems, but in time series the points are correlated! Dropping specific points will mess up the autocorrelation structure.
  - Impute the 0 values (in red) without adjusting the following days (in blue). Why?
- Instead, I would suggest one of the following
  - Consider what may have happened on these 0 days, and adjust accordingly.
  - Drop points of concern and all data points before them too. Anytime we have a finite dataset, we have somehow (arbitrarily?) chosen the start of the time series. You can re-choose this if you feel that
    - Only the "N" most recent datapoints accurately describe the DGP and that
    - No points in your forecast window will come from the process that generated the dropped data

# Data issues? ("Real-world" data is messy ;)





$\nabla Y_t$
Change in Stock Price
Date
5
0
−5
2016
2017
2018
2019
2020
2021

# Tips

- Remember that the end goal is accurate forecasts for $Y_{n+1}, ..., Y_{n+10}$.
- The beginning of the pandemic brought economic uncertainty, which is seen in the higher volatility (variance) during this period. It's unrelated to the mean, and as you've seen, our VST methods don't really affect this form of heteroscedasticity.
- There are tools for modeling/dealing with this, but they're beyond the scope of this course (e.g. GARCH, stochastic volatility models, structural break/regime switching models).
- Personally, I would not drop data points in the middle of my dataset, e.g. 2018 or early 2020, as discussed before.
- Instead, I would suggest one of the following
  - Ignore the change in variance. When defending "stationarity" of your process, you can cite this permission :) Generally, If you estimate the mean well, then your forecasts will be okay. But I know this might feel unsatisfactory for you...
  - Drop points of concern and all data points before them too, as discussed before.