Signal vs. Noise

Jared Fisher

January 23, 2020

Announcements

- Let's look at Homework 0 briefly. Remember: this is not graded! But it will be used to see who on the waitlist/CE is actively participating in class.

Announcements

- Let's look at Homework 0 briefly. Remember: this is not graded! But it will be used to see who on the waitlist/CE is actively participating in class.
- Wifi is down in Evans, so it may not be posted yet.

Announcements

- Let's look at Homework 0 briefly. Remember: this is not graded! But it will be used to see who on the waitlist/CE is actively participating in class.
- Wifi is down in Evans, so it may not be posted yet.
- Homework 0 is due Friday, 1/31, by 8:59am.

Announcements

- Let's look at Homework 0 briefly. Remember: this is not graded! But it will be used to see who on the waitlist/CE is actively participating in class.
- Wifi is down in Evans, so it may not be posted yet.
- Homework 0 is due Friday, 1/31, by 8:59am.
- Must be submitted as a PDF.

Waitlist and Concurrent Students

- Change of plans: I will add anyone on the waitlist/CE to bCourses so you can access materials and Gradescope
- Being on bCourses does not mean you are enrolled in the course

Concurrent Enrollment

- I will begin processing applications when waitlist has emptied
- I will accept the applications of students who have completed all assignments, in the order applications were received
- So! The best way to get in: be up to date on assignments

Accomodations and Schedule Conflicts

- Please let me know of any conflicts or accomodations, religious or otherwise, as soon as possible.

Personal Introduction

- Grew up in Allen, Texas

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves
- University of Texas at Austin

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves
- University of Texas at Austin
  - PhD in Statistics

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves
- University of Texas at Austin
  - PhD in Statistics
    - from the Department of Information, Risk, and Operations Management

Personal Introduction

- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves
- University of Texas at Austin
  - PhD in Statistics
    - from the Department of Information, Risk, and Operations Management
    - Financial time series

Personal Introduction

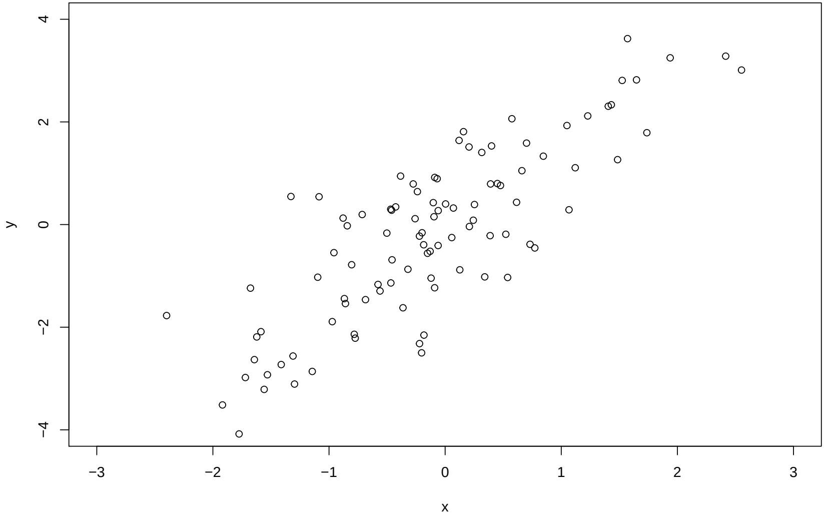
- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves
- University of Texas at Austin
  - PhD in Statistics
    - from the Department of Information, Risk, and Operations Management
    - Financial time series
- 3 kids

Personal Introduction

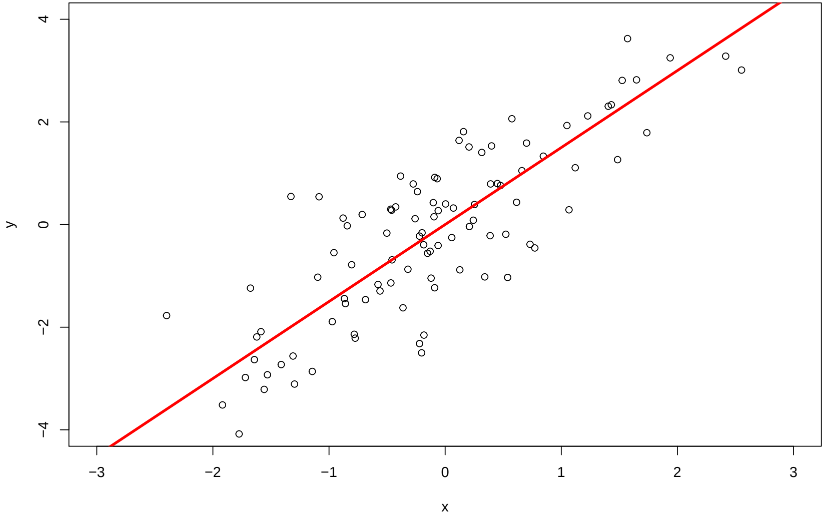
- Grew up in Allen, Texas
- Brigham Young University
  - BS in Statistics
  - MS in Statistics
    - Baseball performance curves
- University of Texas at Austin
  - PhD in Statistics
    - from the Department of Information, Risk, and Operations Management
    - Financial time series
- 3 kids
  - Ages 6, 3, and 0

Parallels between Ideas behind Time Series and Regression

Regression



Regression



Regression = Signal + Noise

- $y_i = f(\mathbf{x_i}) + \epsilon_\mathbf{i}$

Regression $=$ Signal $+$ Noise

- $y_i = f(\mathbf{x_i}) + \epsilon_{\mathbf{i}}$
- $f(\mathbf{x_i}) \Rightarrow$ "Signal"

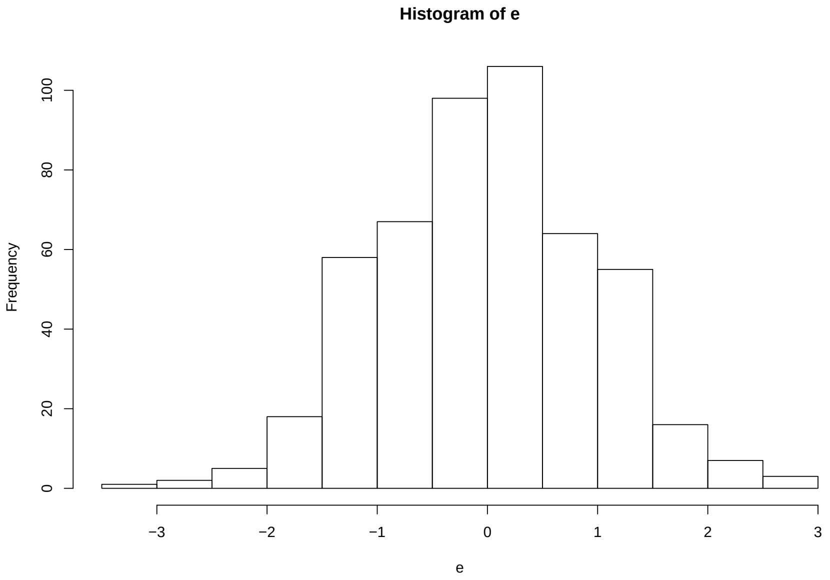
Regression = Signal + Noise

- $y_i = f(\mathbf{x_i}) + \epsilon_\mathbf{i}$
- $f(\mathbf{x_i}) \Rightarrow$ "Signal"
- $\epsilon_i \Rightarrow$ "Noise"

Regression = Signal + Noise

- $y_i = f(\mathbf{x_i}) + \epsilon_\mathbf{i}$
- $f(\mathbf{x_i}) \Rightarrow$ "Signal"
- $\epsilon_i \Rightarrow$ "Noise"
  - Often $\epsilon_i \sim N(0, \sigma^2)$

Gaussian Errors

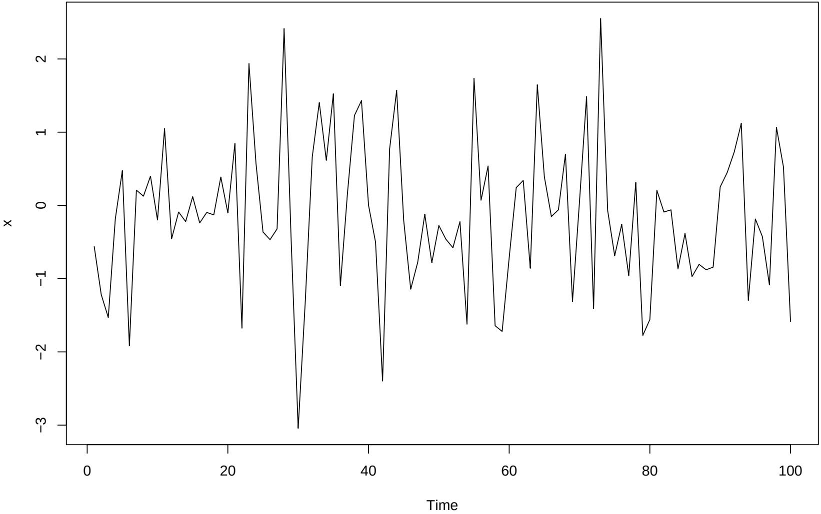


But. . .

- in time series we don't have n subjects at 1 point in time, we have 1 subject at n points in time. . .
- So, instead think of the model as $x_t = f(\mathbf{x}) + \epsilon_{\mathbf{t}}$, and now view the ϵ's over time
- Or, with no signal: $x_t = \epsilon_t$

Gaussian Errors, over time

```
plot.ts(x)
```



We call this "Gaussian noise"

Definitions (TSA4e Example 1.8)

Random variables $X_1, \ldots, X_n$ will be denoted as

- White noise: if they have mean zero, variance σ^2, and are uncorrelated

Definitions (TSA4e Example 1.8)

Random variables $X_1, ..., X_n$ will be denoted as

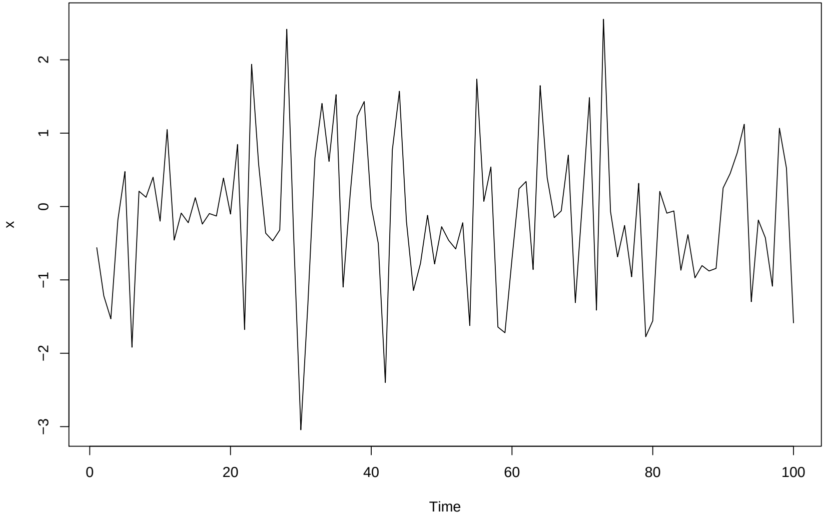
- White noise: if they have mean zero, variance σ^2, and are uncorrelated
- IID noise: if they are white noise AND are independent and identically distributed (IID).

Definitions (TSA4e Example 1.8)

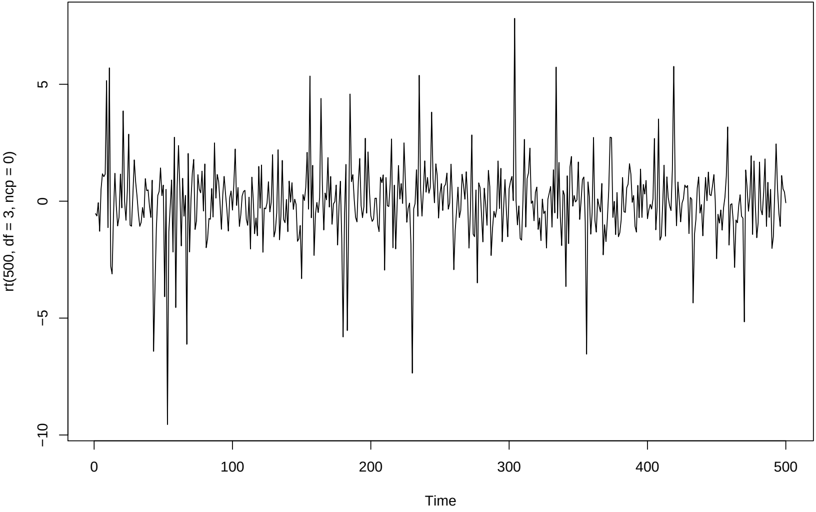
Random variables $X_1, \ldots, X_n$ will be denoted as

- White noise: if they have mean zero, variance σ^2, and are uncorrelated
- IID noise: if they are white noise AND are independent and identically distributed (IID).
- Gaussian [white] noise: if they are IID noise AND are normally distributed, $X_i \sim N(0, \sigma^2)$

Gaussian Noise



IID Noise (T distribution)



But wait!

- Didn't we say that this class is about things that are correlated over time?

But wait!

- Didn't we say that this class is about things that are correlated over time?
- White noise has no varying structure over time, so for most time series data it's not a good model

But wait!

- Didn't we say that this class is about things that are correlated over time?
- White noise has no varying structure over time, so for most time series data it's not a good model
- BUT it is the basis for many time series models

But wait!

- Didn't we say that this class is about things that are correlated over time?
- White noise has no varying structure over time, so for most time series data it's not a good model
- BUT it is the basis for many time series models
- So how do we check to see if white noise is an appropriate model?

Tools we'll need

- Autocovariance (Definition 1.2):

$$\begin{aligned}\gamma_x(s,t) &= cov(X_s, X_t)\\ &= E\left[(X_s - E[X_s])(X_t - E[X_t])\right]\end{aligned}$$

Tools we'll need

- Autocovariance (Definition 1.2):

$$\begin{aligned}\gamma_x(s,t) &= cov(X_s, X_t)\\ &= E\left[(X_s - E[X_s])(X_t - E[X_t])\right]\end{aligned}$$

- [mental aside: let $s > t$ and $h = s - t$. h is the number of "lags"]

Tools we'll need

- Autocovariance (Definition 1.2):

$$\begin{aligned}\gamma_x(s,t) &= cov(X_s, X_t)\\ &= E\left[(X_s - E[X_s])(X_t - E[X_t])\right]\end{aligned}$$

- [mental aside: let $s > t$ and $h = s - t$. h is the number of "lags"]
- Sample autocovariance (Definition 1.14):

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x})$$

Tools we'll need

- Autocorrelation function "ACF" (Definition 1.3):

$$\rho(s,t) = \frac{\gamma_x(s,t)}{\sqrt{\gamma_x(s,s)\gamma_x(t,t)}} = \frac{cov(X_s, X_t)}{\sqrt{var(X_s)var(X_t)}}$$

Tools we'll need

- Autocorrelation function "ACF" (Definition 1.3):

$$\rho(s,t) = \frac{\gamma_x(s,t)}{\sqrt{\gamma_x(s,s)\gamma_x(t,t)}} = \frac{cov(X_s, X_t)}{\sqrt{var(X_s)var(X_t)}}$$

- Sample autocorrelation (Definition 1.15):

$$r_h = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-h}(x_t - \bar{x})(x_{t+h} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

A few notes:

- The book uses $\hat{\rho}$. I'll accept either r or $\hat{\rho}$ for sample correlations.

A few notes:

- The book uses $\hat{\rho}$. I'll accept either r or $\hat{\rho}$ for sample correlations.
- Denominator is the sample variance of X_t, not product of the sample standard deviations of X_t and X_{t+h}, i.e. $\sqrt{\hat{var}(X_{t+h})\hat{var}(X_t)}$

A few notes:

- The book uses $\hat{\rho}$. I'll accept either r or $\hat{\rho}$ for sample correlations.
- Denominator is the sample variance of X_t, not product of the sample standard deviations of X_t and X_{t+h}, i.e. $\sqrt{\hat{var}(X_{t+h})\hat{var}(X_t)}$
- We're assuming something to the effect of $\hat{var}(X_{t+h}) \overset{?}{\approx} \hat{var}(X_t)$

A few notes:

- The book uses $\hat{\rho}$. I'll accept either r or $\hat{\rho}$ for sample correlations.
- Denominator is the sample variance of X_t, not product of the sample standard deviations of X_t and X_{t+h}, i.e. $\sqrt{\hat{var}(X_{t+h})\hat{var}(X_t)}$
- We're assuming something to the effect of $\hat{var}(X_{t+h}) \overset{?}{\approx} \hat{var}(X_t)$
- (Not a big deal right now, but a peculiar piece with the estimator)

Properties of White noise

- $E(X_t) = 0$
- $var(X_t) = \sigma^2$ (constant)
- $\rho(s, t) = 0$ for all $s \neq t$
- $\rho(t, t) = 1$ (by obvious)

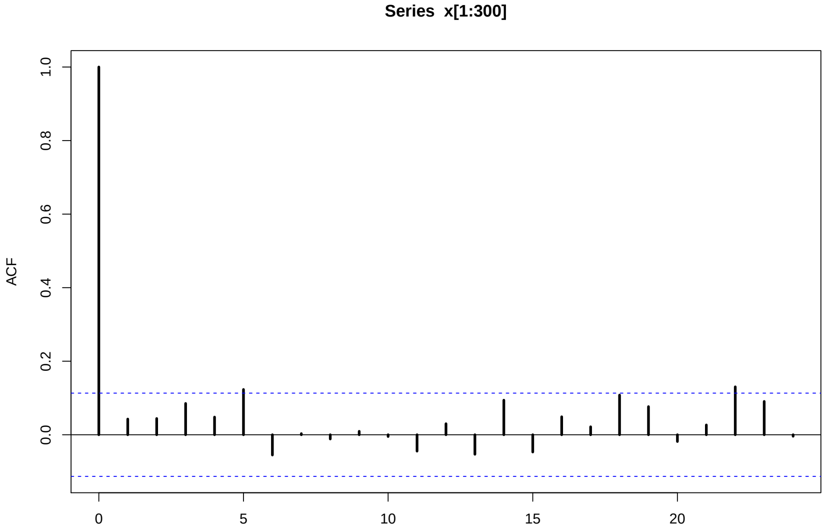
- How do we check if white noise is a reasonable model for a time series?

Properties of White noise

- $E(X_t) = 0$
- $var(X_t) = \sigma^2$ (constant)
- $\rho(s, t) = 0$ for all $s \neq t$
- $\rho(t, t) = 1$ (by obvious)
- How do we check if white noise is a reasonable model for a time series?
- Question we'll ask ourselves: for large enough n, $r_k \approx \rho(k) = 0$ for all $k \neq 0$?

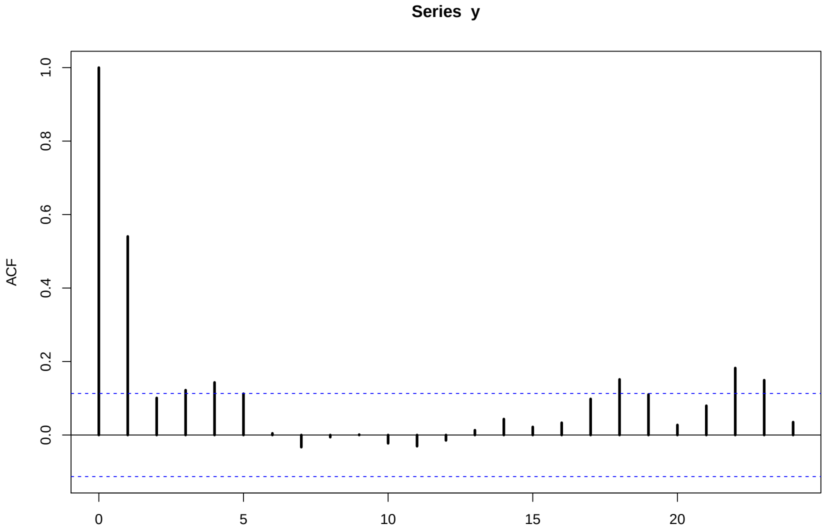
ACF plot

```r
x = rnorm(301)
acf(x[1:300],lwd=3)
```



ACF plot - simple Moving Average

```
y = .5*(x[1:300] + x[2:301])
acf(y,lwd=3)
```



CI for Sample Correlations

- Wouldn't it be great if those dashed blue lines were the appropriate confidence interval?

Simplified Theorem A.7 (see Property 1.2)

- Under general conditions, if x_t is white noise, then for n large, and with arbitrary but fixed H, then the sample autocorrelations

$$r_1, ..., r_H \overset{iid}{\sim} N(0, 1/n)$$

- In other words

$$\sqrt{n}\begin{pmatrix} r_1 \\ \vdots \\ r_H \end{pmatrix} \rightarrow N(0, I) \quad \text{as } n \rightarrow \infty$$

- Key takeaway: $var(r_h) = 1/n$ (Equation 1.38)

Confidence Interval

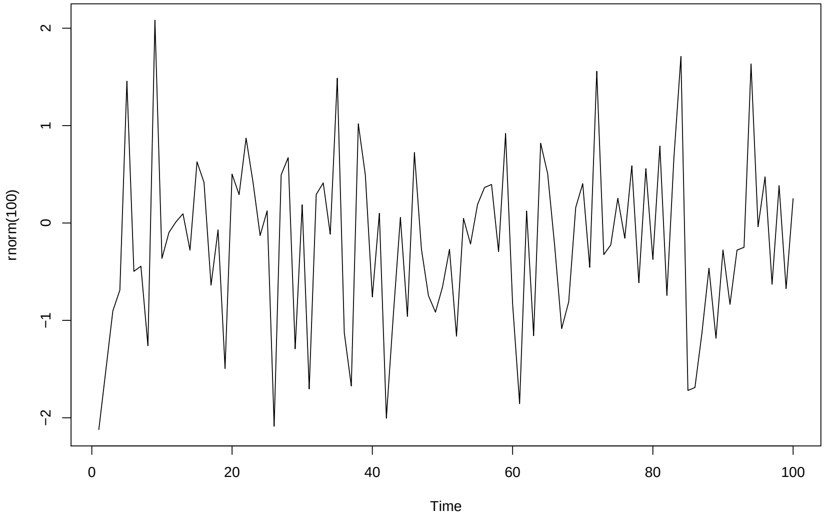
- For white noise $r_1, ..., r_H \overset{iid}{\sim} N(0, 1/n)$, for

  $$P\left(|r_h| > 1.96n^{-\frac{1}{2}}\right) \approx P\left(|N(0,1)| > 1.96\right) = 5\%$$

- So for $n = 100$, $1.96n^{-\frac{1}{2}} = 1.96/\sqrt{100} = .196$

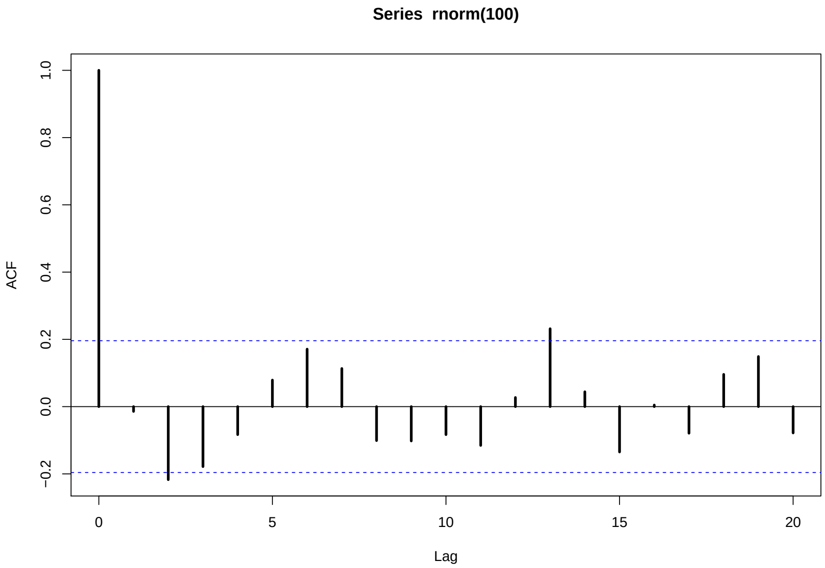
Gaussian Noise

```
plot.ts(rnorm(100))
```



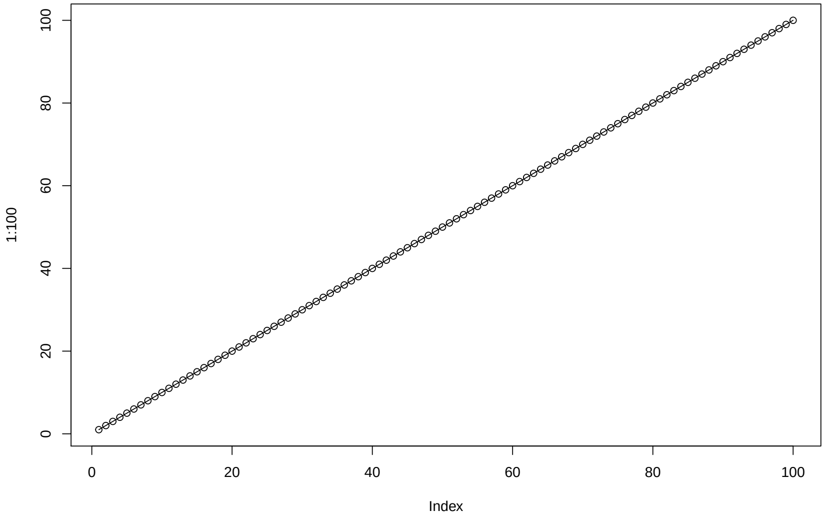
ACF plot - Dashes at .196?

```
acf(rnorm(100),lwd=3)
```



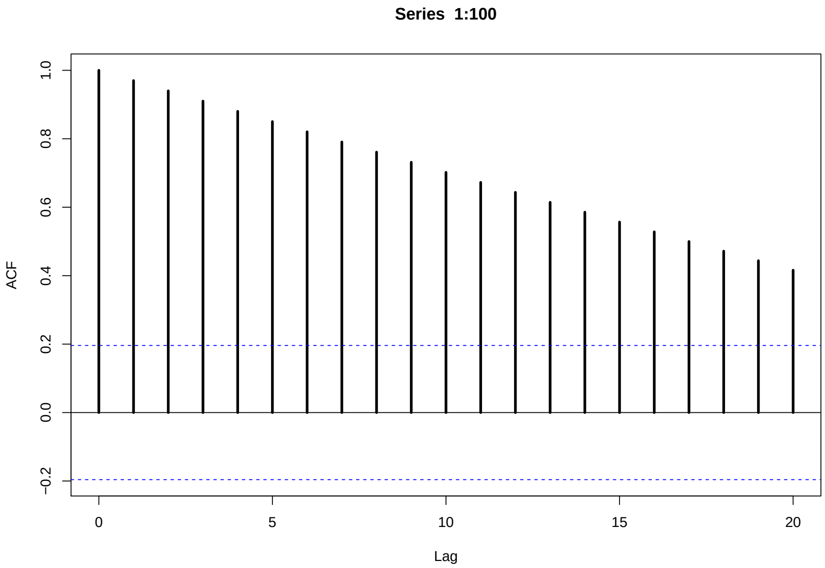
Straight line

```
plot(1:100,type='o')
```



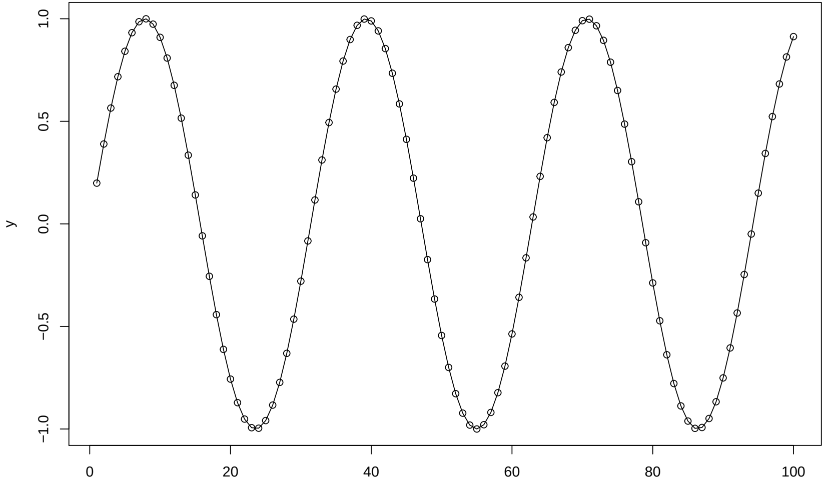
ACF plot - Dashes at .196?

```
acf(1:100,lwd=3)
```



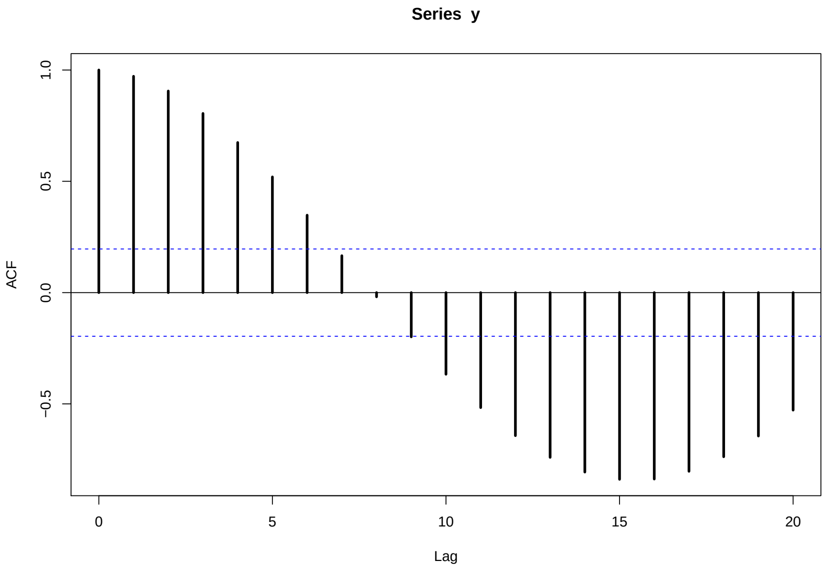
Sine function

```
x = 1:100
y = sin(x/5)
plot(x,y,type='o')
```



ACF plot - Dashes at .196?

```
acf(y,lwd=3)
```



Final Thought

The ACF plot is called a “Correlogram”. We use the correlogram of the ACF to diagnose whether white noise is a reasonable model for a time series.