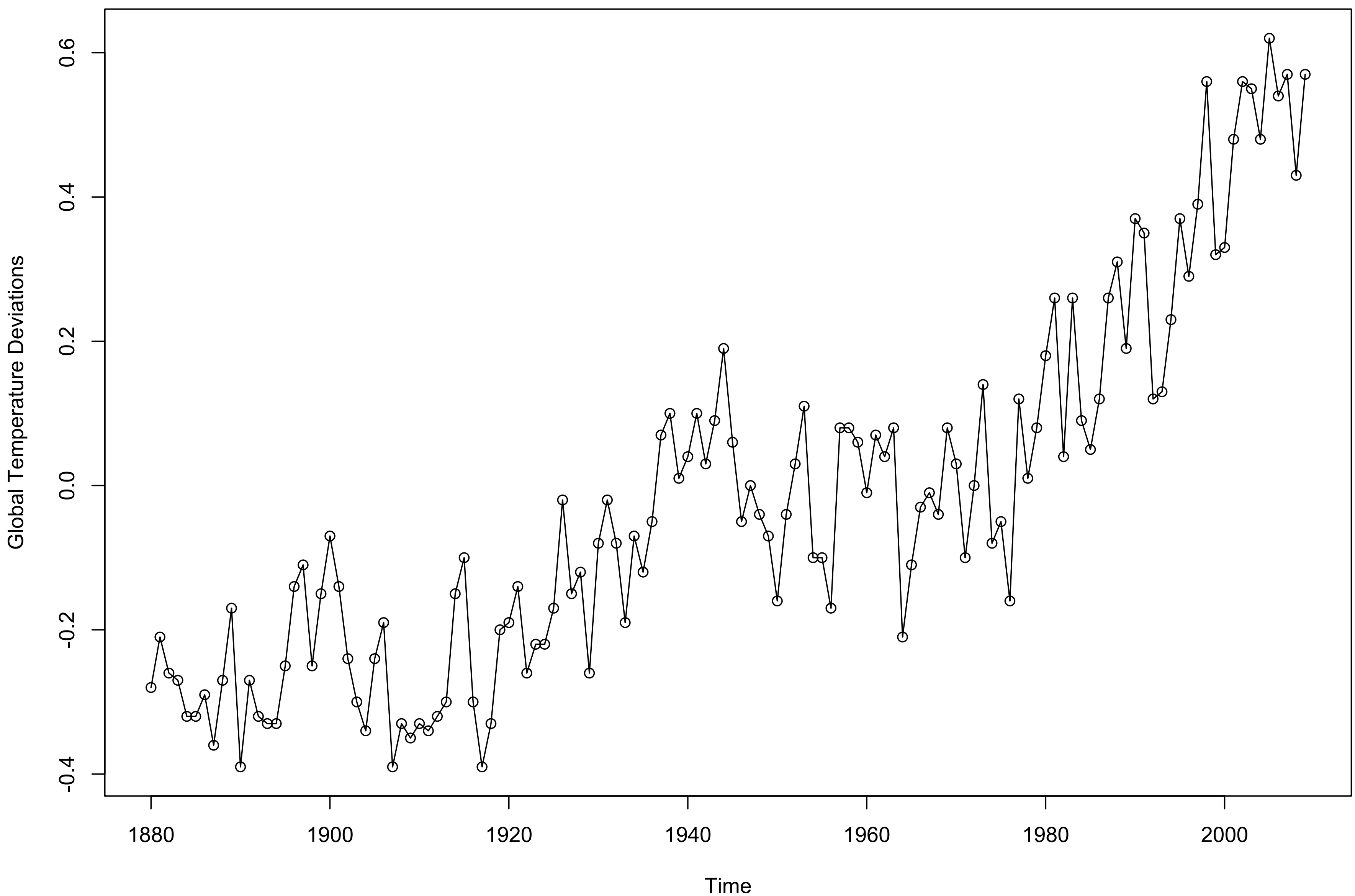
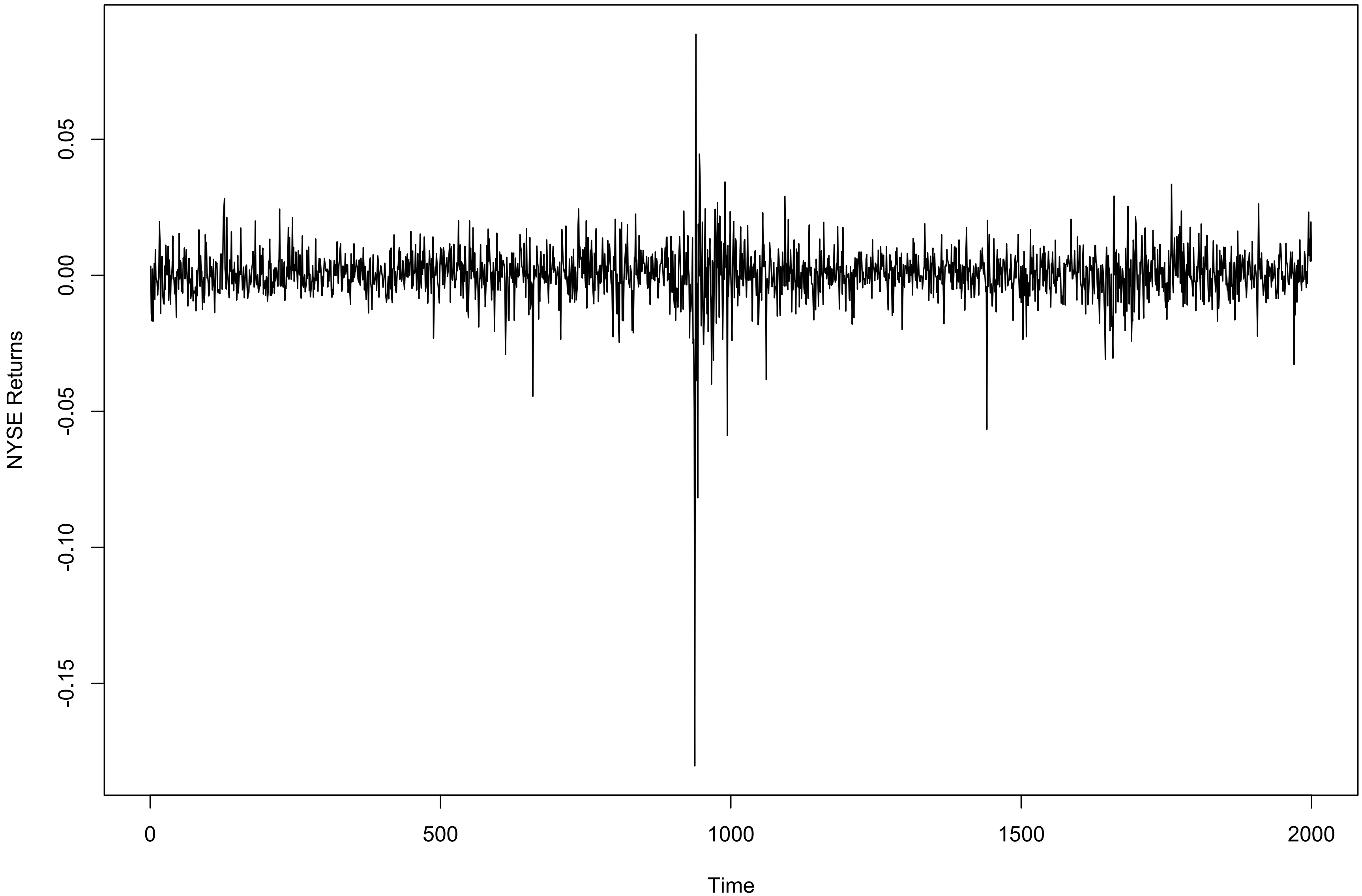
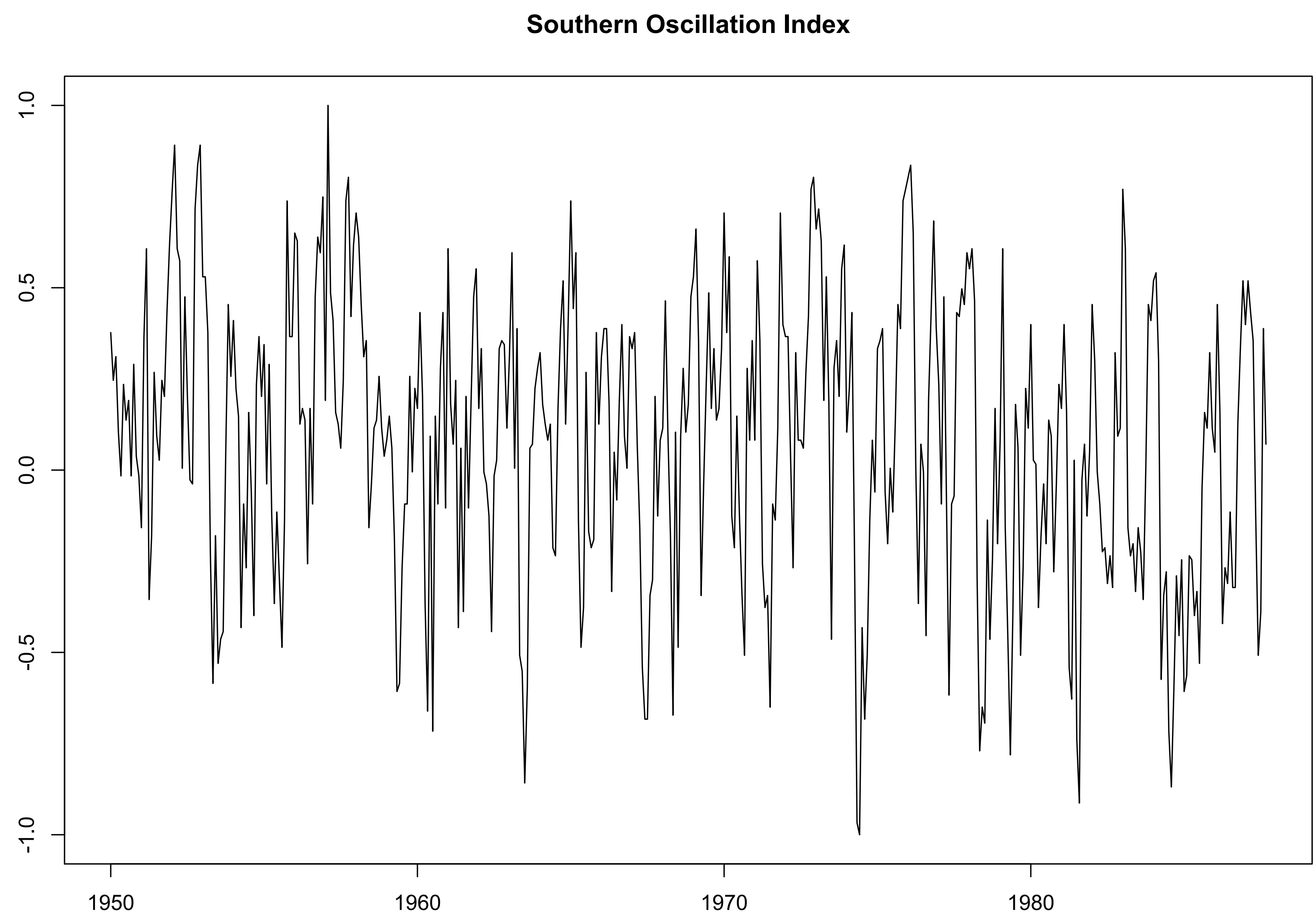
# An Overview of Time Series Analysis




Global Temperatures since 1880




Returns to NYSE Feb 2 1984 – Dec. 1991




Southern Oscillation Index (SOI): Index which measures sea level pressure – typically correspond to intensity of El Nino an La Nina.

# Common Problems with Times Series Data

1. Temporal Correlation
2. Temporal Trends
3. Seasonal Variation
4. Heteroskedasticity (particularly for stock data)

# Important Definitions for Time Series Data

Mean Function:

$$\mu_t = \mathbb{E}(x_t)$$

Autocovariance Function:

$$\gamma(t_1, t_2) = \mathrm{Cov}(x_{t_1}, x_{t_2})$$

Autocorrelation function:

$$\rho(t_1, t_2) = \mathrm{Corr}(x_{t_1}, x_{t_2})$$

# Important Definitions for Time Series Data

Stationary Time Series: Distribution of a time series is invariant to shifts in time,

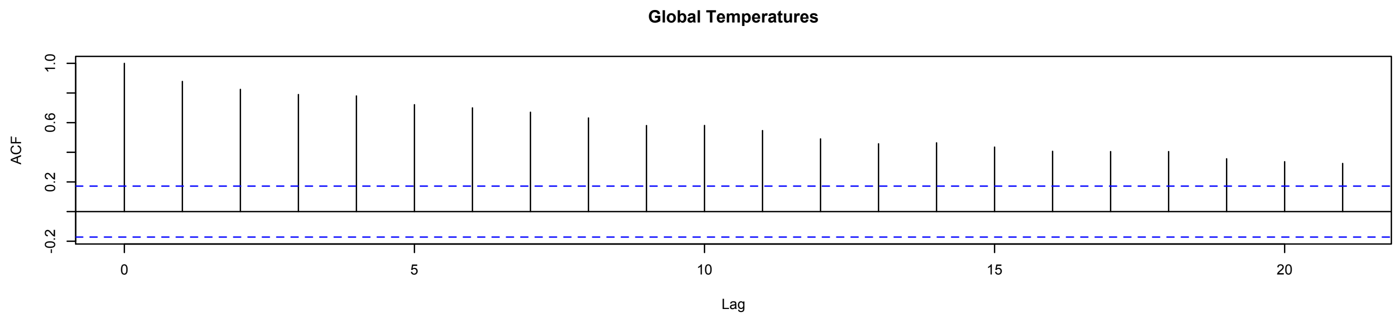
$$\text{Prob}(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \ldots, x_{t_T} \leq c_T) = \text{Prob}(x_{t_1+h} \leq c_1, x_{t_2+h} \leq c_2, \ldots, x_{t_T+h} \leq c_T)$$

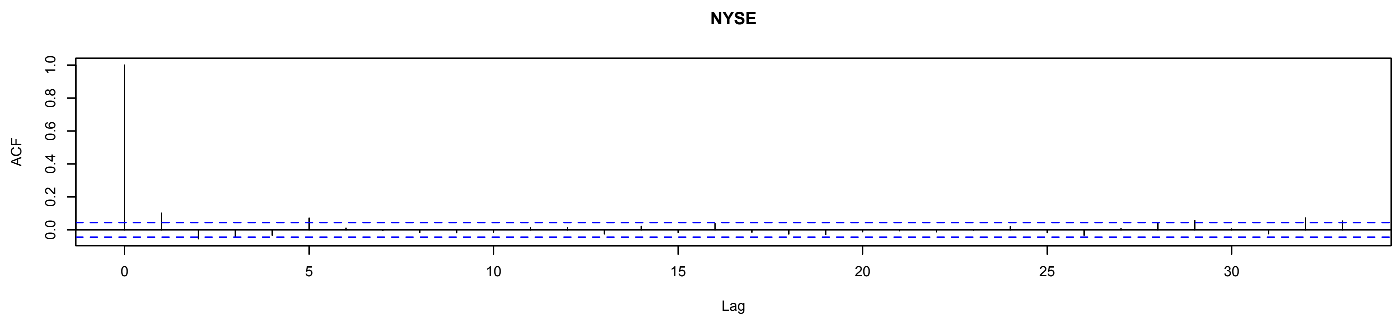
Weakly Stationary Time Series:

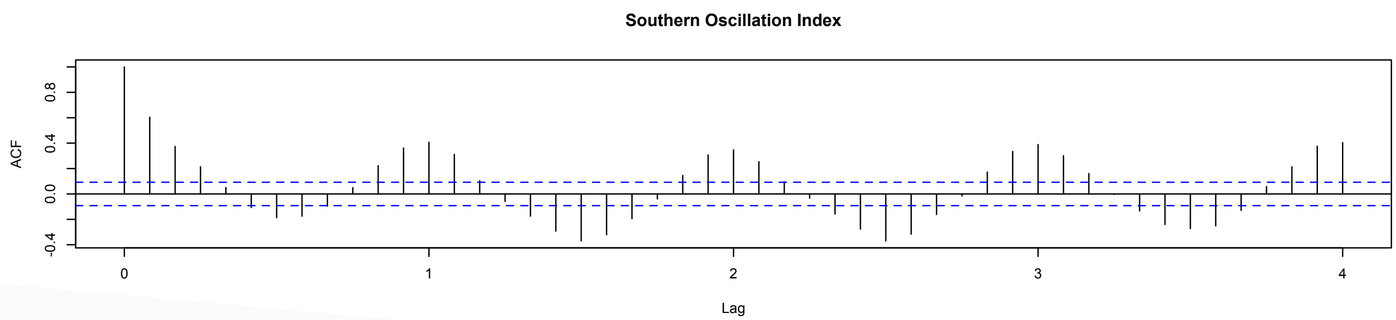
1. Mean does not change in time: $\mu_t = \mu$
2. Covariance only depends on separation:

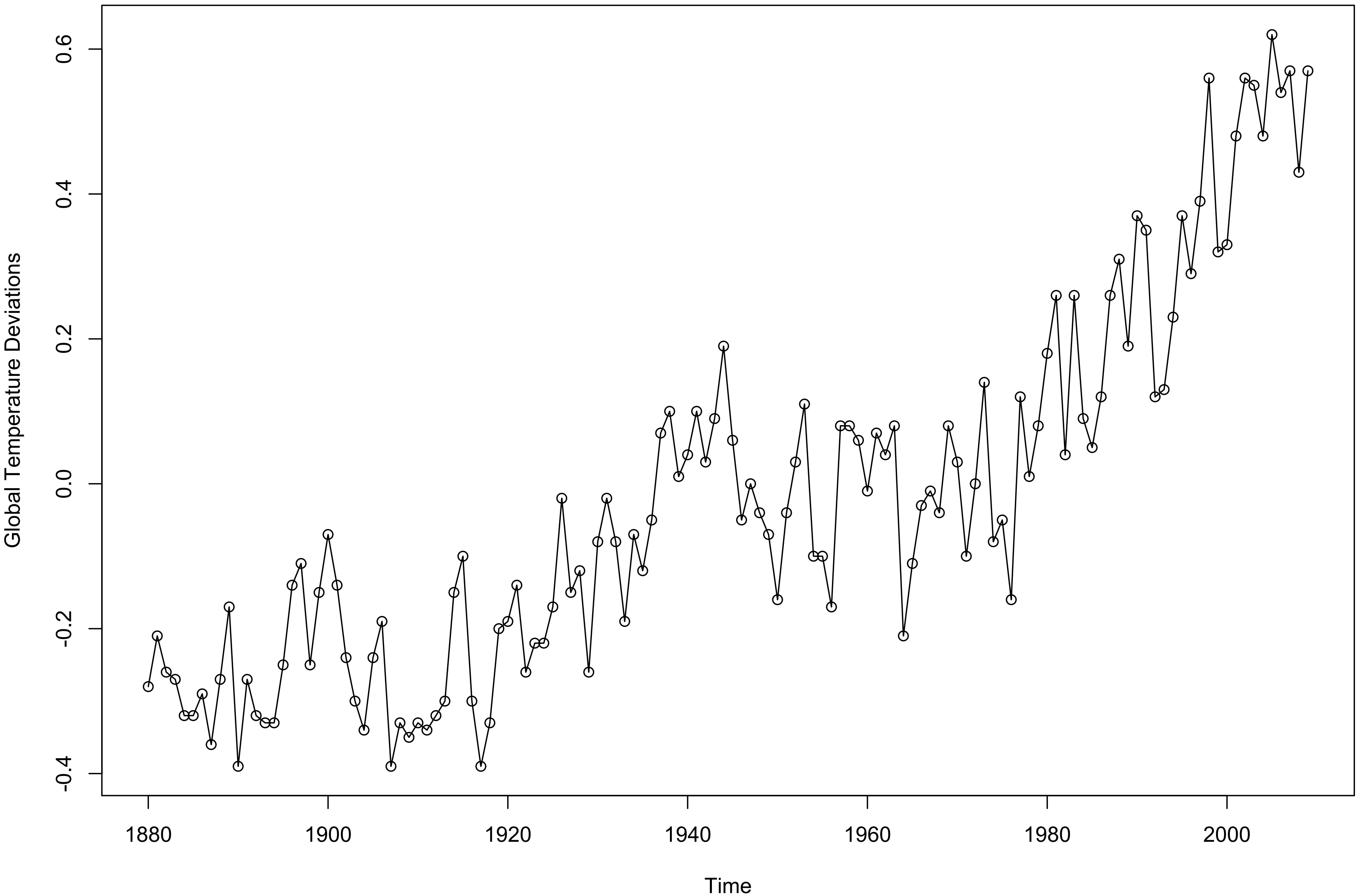
$$\gamma(t_1, t_2) = \gamma(t_1 - t_2) = \gamma(h)$$

# Examples Autocorrelation Functions



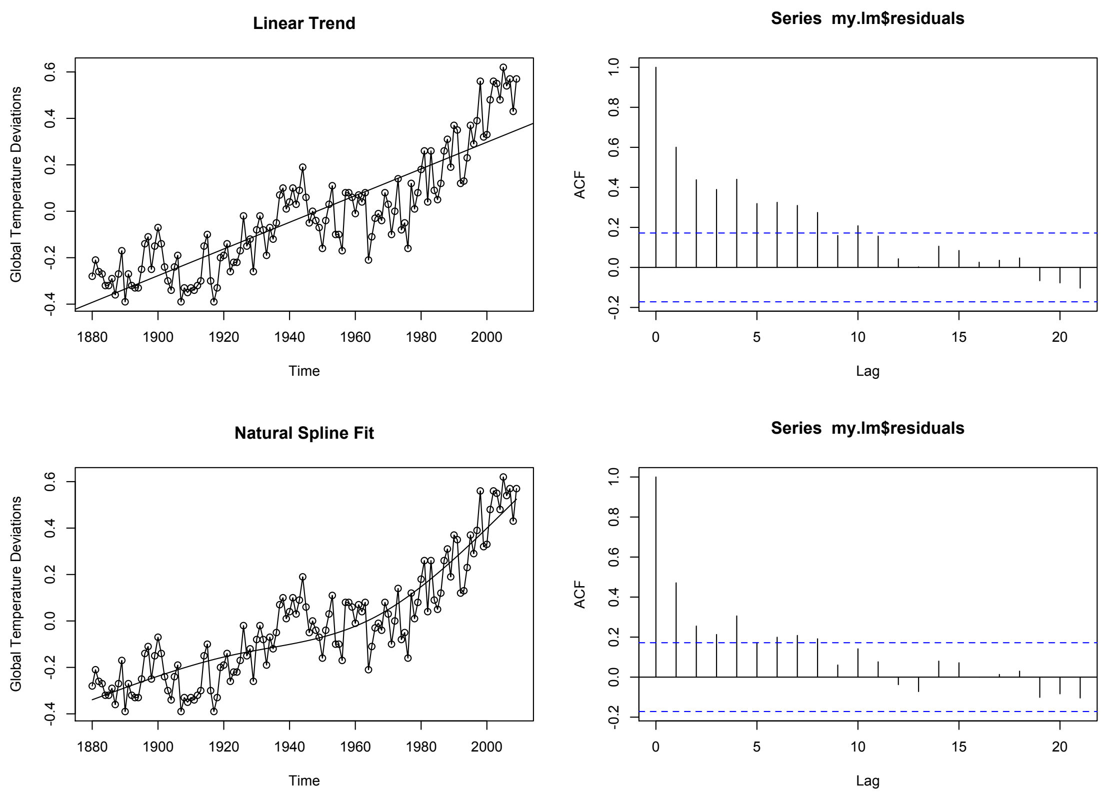



What would you do with the data: $x_1, \ldots, x_T$?

# Regression for Time Series

# Regression for Time Series

Big Idea: Treat time (t) as a “covariate” in a regression model.

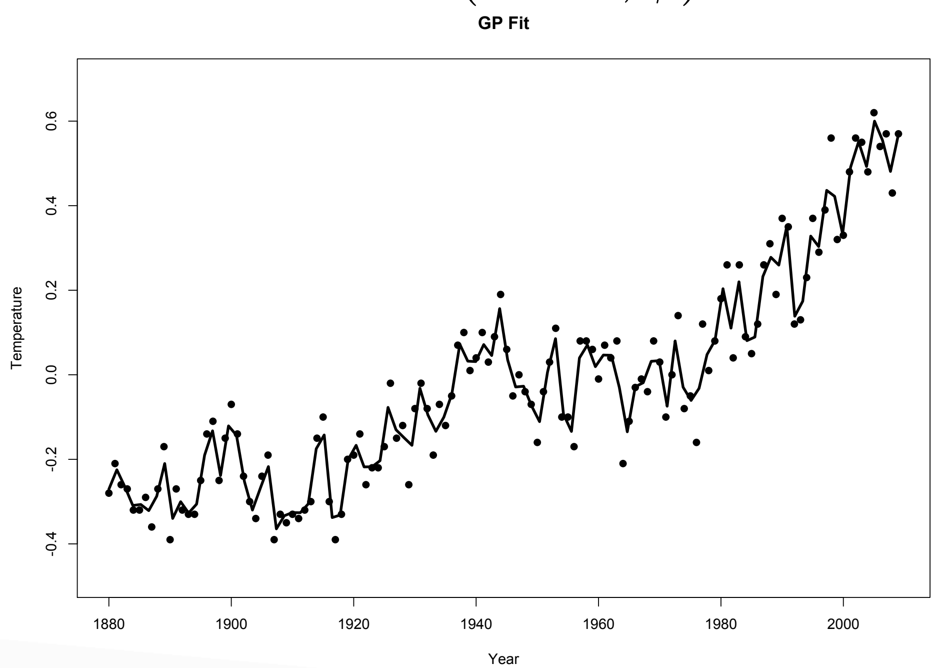


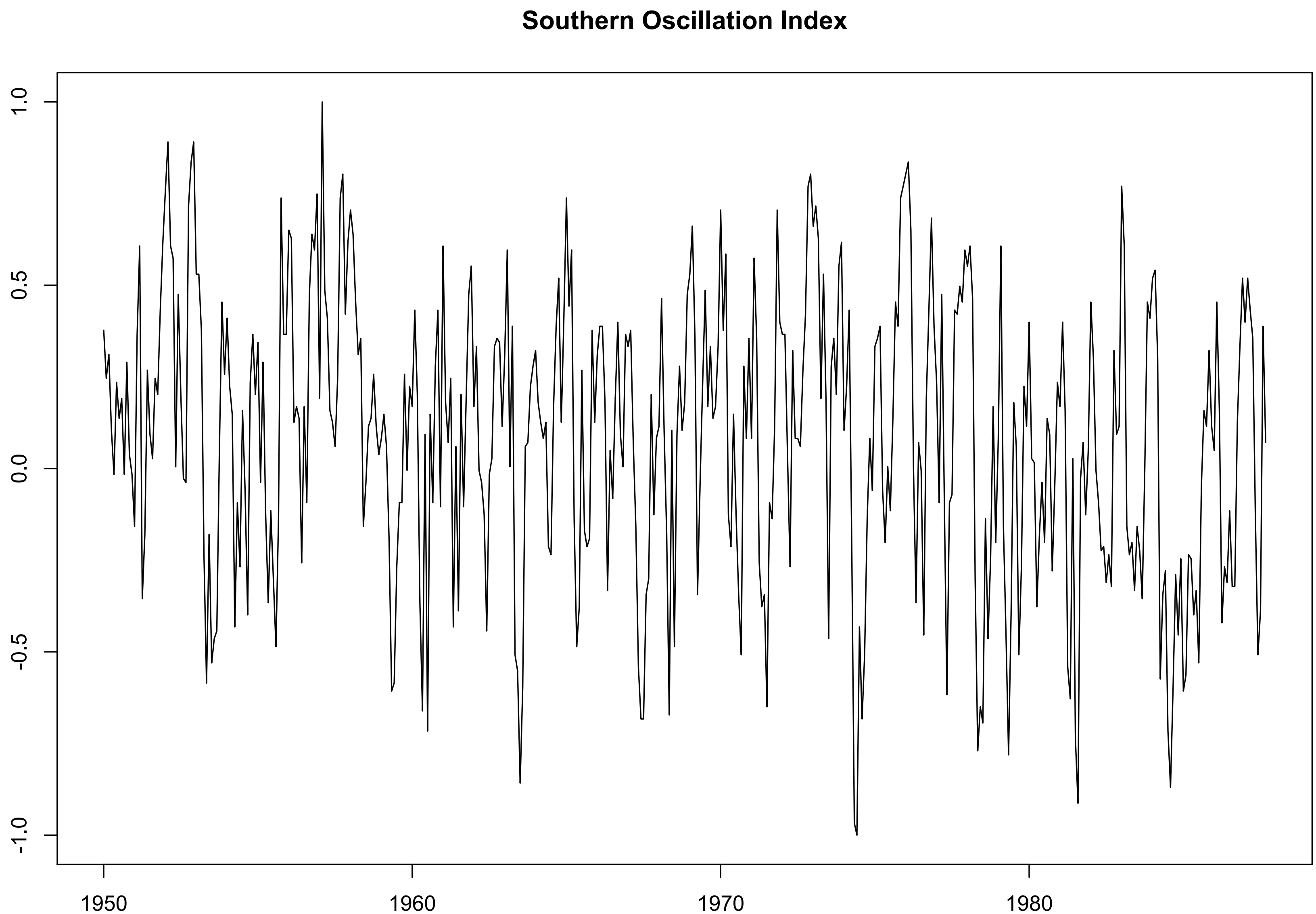
# Gaussian Process Regression for Time Series

Big Idea: Treat $x_{t_1}, \ldots, x_{t_T}$ as a realization of a GP over the time domain $t_1, \ldots, t_T$.

$$(x_{t_1}, \ldots, x_{t_T})' \sim \mathcal{N}(\beta_0 + \text{ns}(t_1, \ldots, t_T, \text{df} = 5)\boldsymbol{\beta}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\Sigma} = \text{Matern}(\nu = 2, \phi)$$





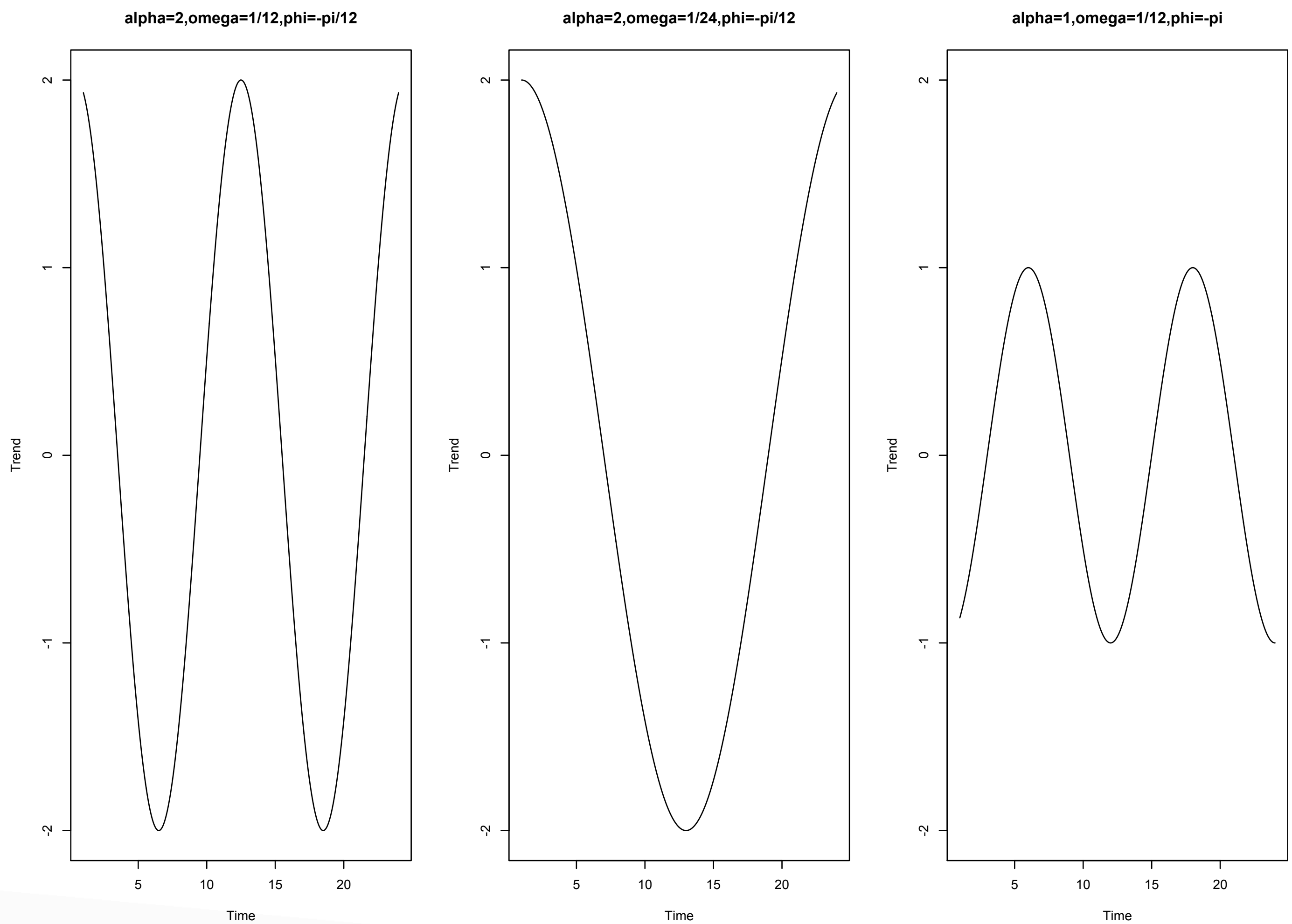

How would you account for the seasonality in the SOI data?

# Regression for Seasonal Time Series

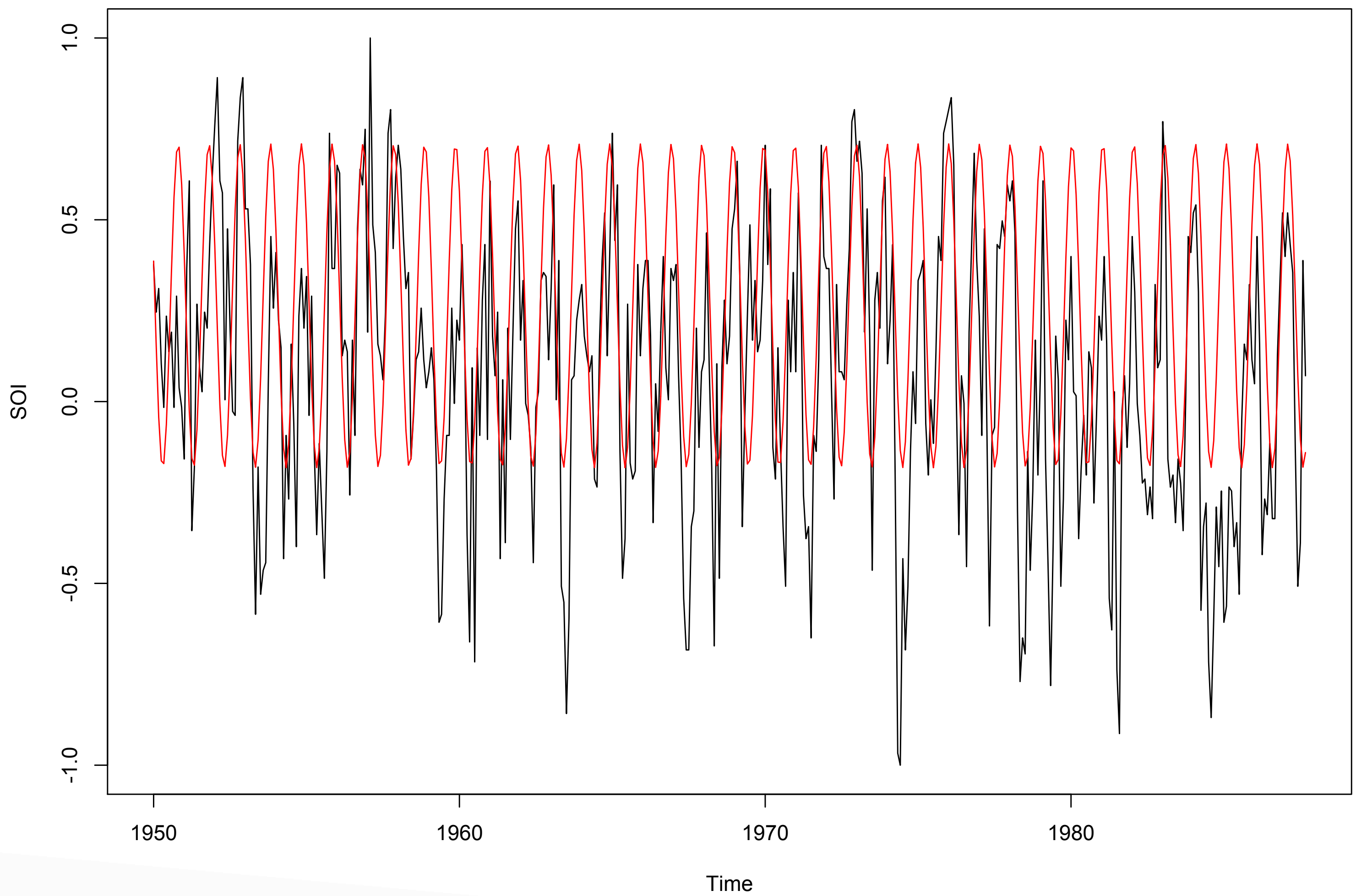
Big Idea: Use a seasonal function (like a cosine)

$$
\begin{aligned}
x_t &= \alpha \cos(2\pi\omega t + \phi) + \epsilon_t \\
&= \alpha \cos(\phi) \cos(2\pi\omega t) + \alpha \sin(\phi) \sin(2\pi\omega t) + \epsilon_t \\
&= \beta_1 \cos(2\pi\omega t) + \beta_2 \sin(2\pi\omega t) + \epsilon_t \\
\alpha &= \text{Amplitude: Seasonal highs/lows} \\
\omega &= \text{Frequency in cycles per time period} \\
\phi &= \text{Phase shift:}
\end{aligned}
$$

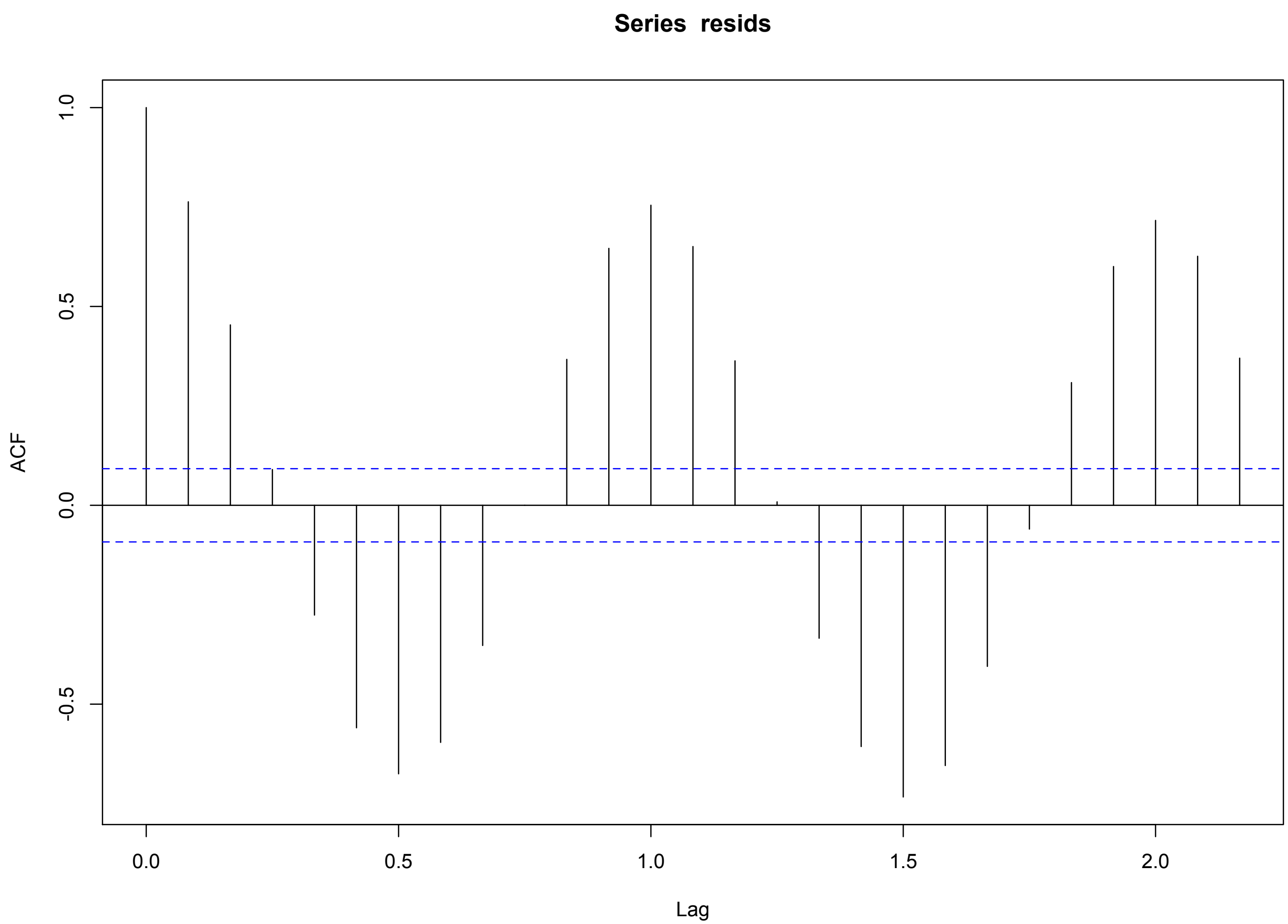
# Regression for Seasonal Time Series



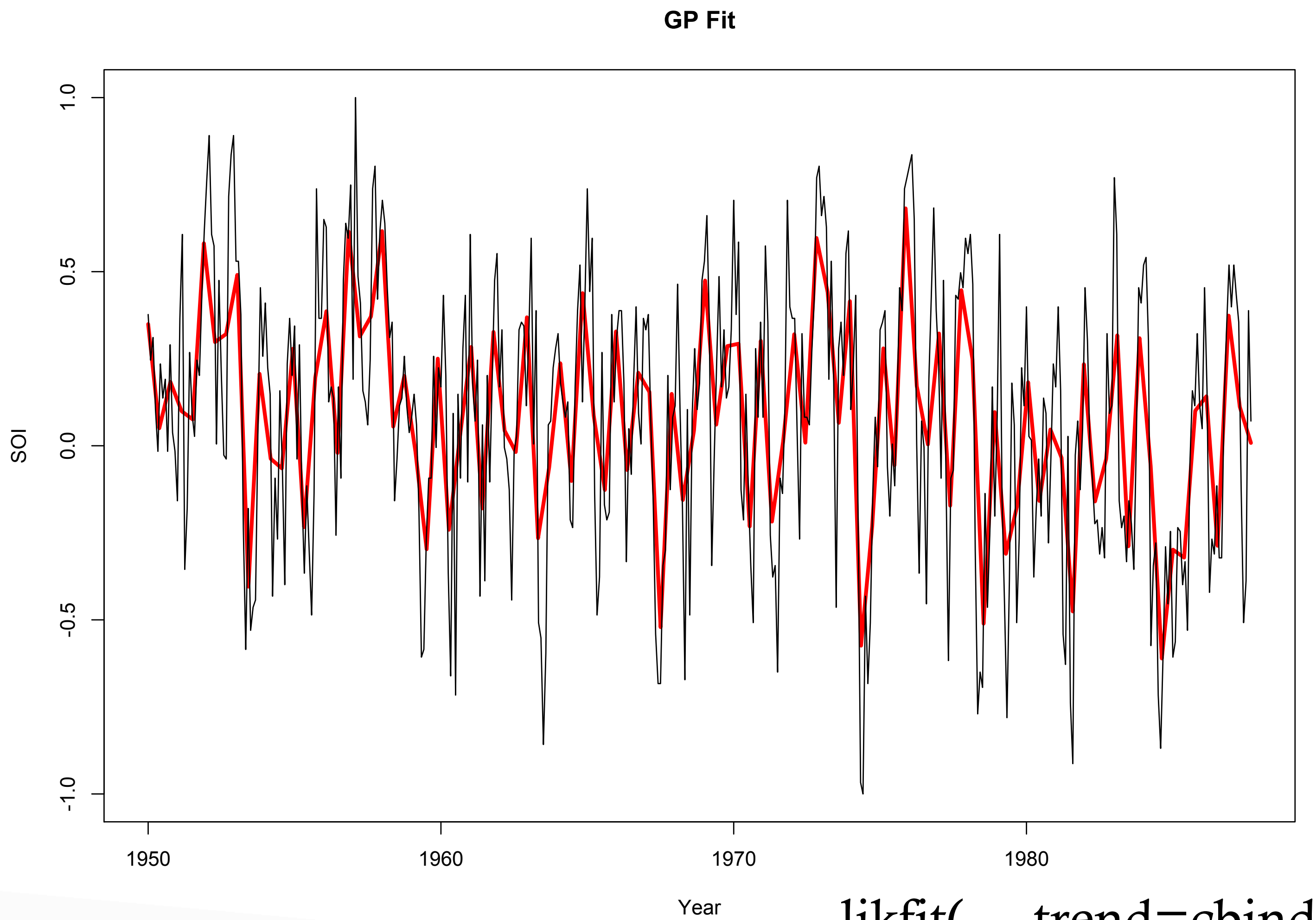
# Regression for Seasonal Time Series



# Regression for Seasonal Time Series



# Regression for Seasonal Time Series




likfit(…,trend=cbind(cos(),sin()))

# ARIMA Models

# Autoregressive Processes

Big Idea: Treat past values as predictors

$$\mathrm{AR}(p): x_t = \mu + \phi_1(x_{t-1} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + \epsilon_t$$

$$= \alpha + \boxed{\sum_{l=1}^{p} \phi_l x_{t-l}} + \epsilon_t$$

Multicollinearity!!

$$\alpha = \mu(1 - \sum_{l=1}^{p} \phi_l), \;\; \epsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

---

$$\mathrm{AR}(1): x_t = \mu + \phi_1(x_{t-1} - \mu) + \epsilon_t$$

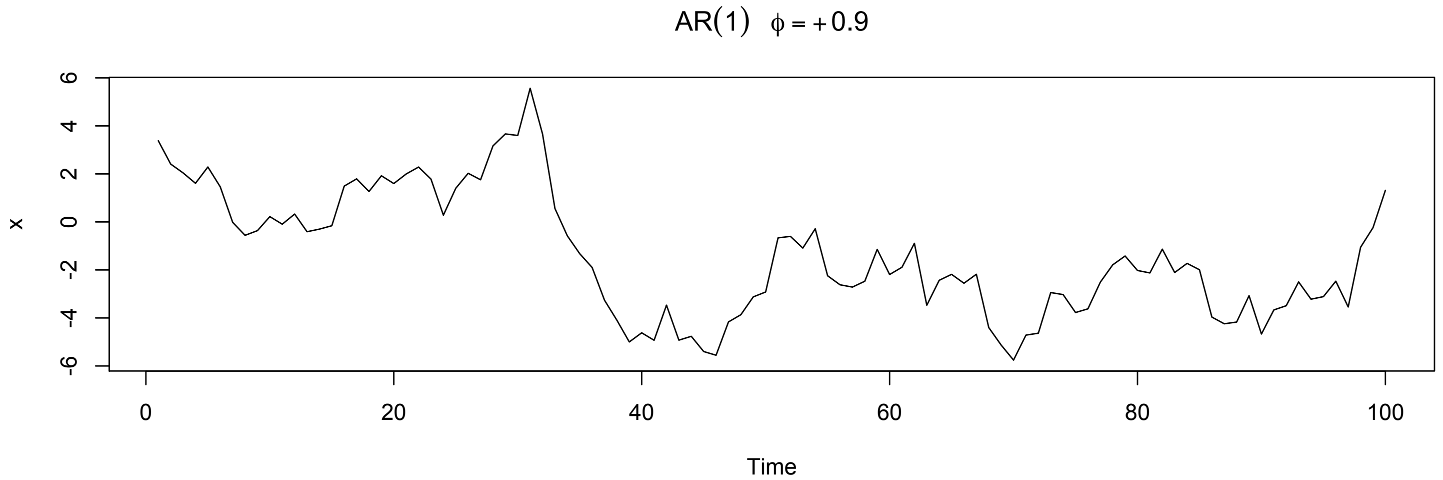
$$\mathbb{E}(\mathrm{AR}(1)) = \mu$$

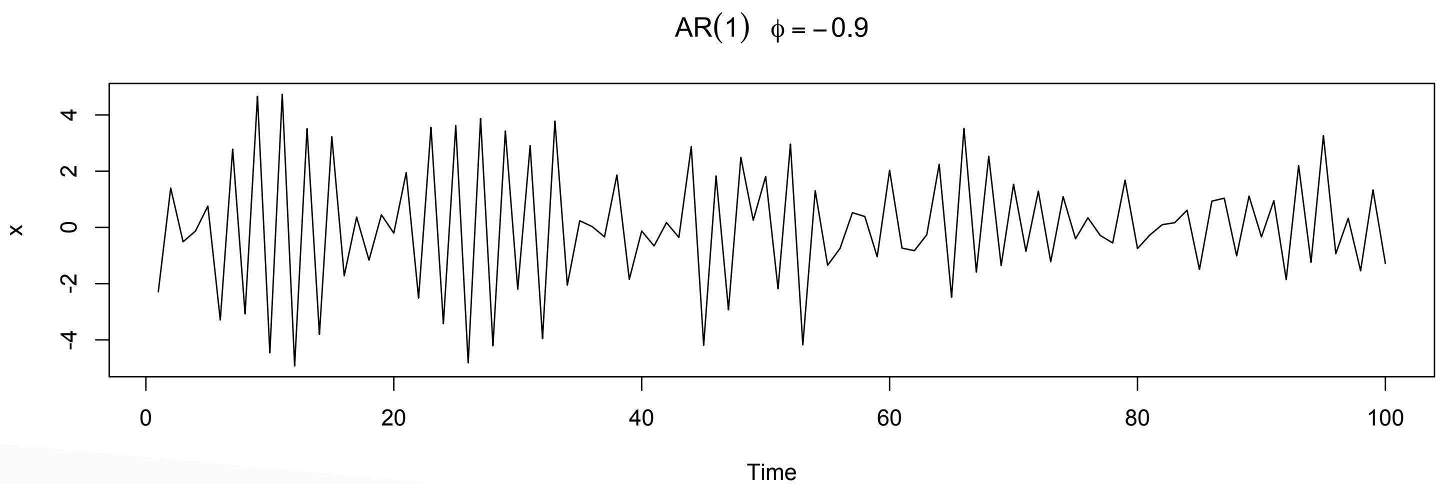
$$\rho(x_t, x_{t-l}) = \phi_1^l \rightarrow |\phi_1| < 1$$

$$\phi_1 < 0 \Rightarrow \text{Negative Correlation}$$

Side Note: You can get similar behavior using the Matern covariance function in a Gaussian process.

# Autoregressive Processes



# Moving Average Processes

Big Idea: Smooth-out Random Shocks

$$\text{MA}(q): x_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t$$

$$\epsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

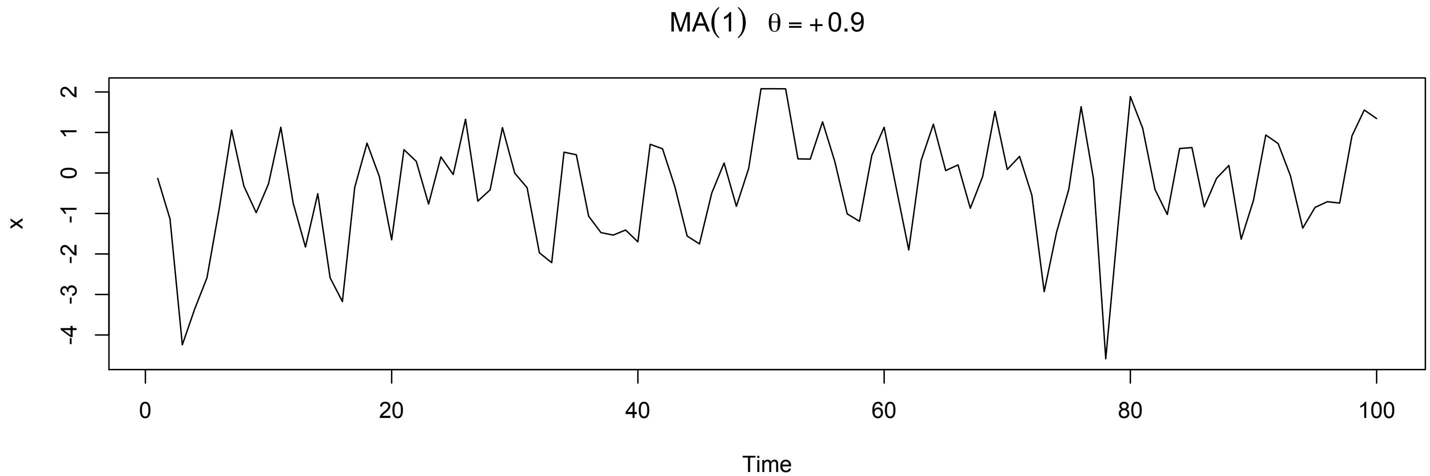
$$\text{MA}(1): x_t = \mu + \theta_1 \epsilon_{t-1} + \epsilon_t$$

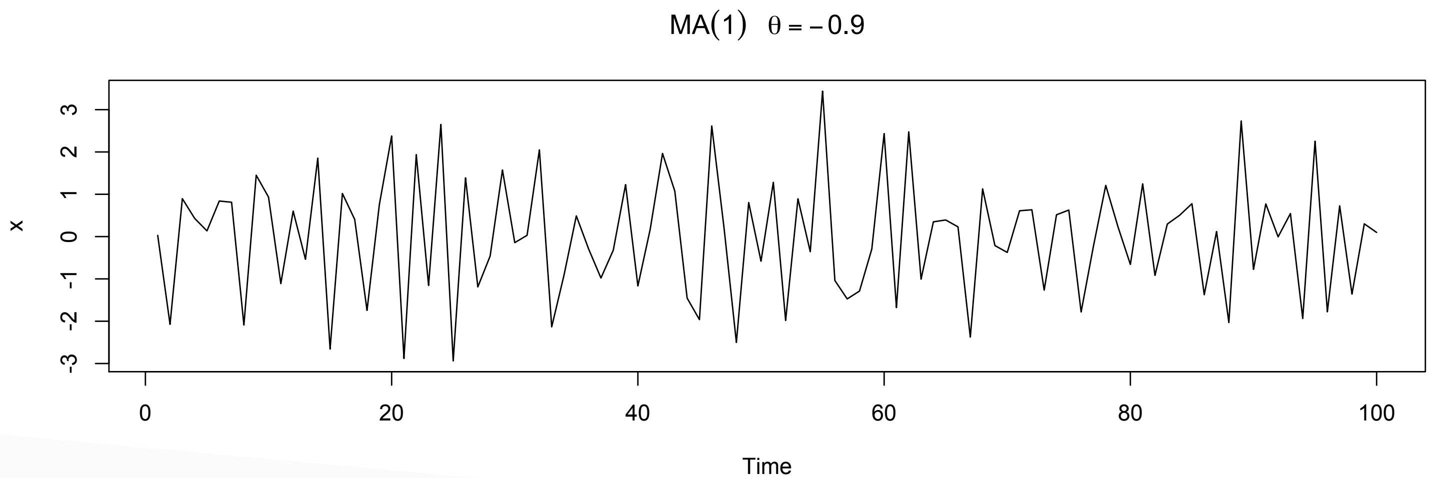
$$\mathbb{E}(\text{MA}(1)) = \mu$$

$$\rho(x_t, x_{t-l}) = \begin{cases} \frac{\theta_1}{1+\theta_1^2} & \text{if } l = 1 \\ 0 & \text{otherwise} \end{cases}$$

Side Note: You can get similar behavior using a tapered covariance function in a Gaussian process (e.g. Wendland).

# Moving Average Processes



# ARMA Processes

Big Idea: Regress on past values of the process but have correlated error terms:

$$\text{ARMA}(p,q): x_t = \mu + \sum_{l_1=1}^{p} \phi_{l_1}(x_{t-l_1} - \mu) + \sum_{l_2=1}^{q} \theta_{l_2}\epsilon_{t-l_2} + \epsilon_t$$

$$x_t = \beta_0 + \sum_{l=1}^{p} \beta_l x_{t-1} + \epsilon_t^{\star}$$

$$\epsilon_t^{\star} = \sum_{l_2=1}^{q} \theta_{l_2}\epsilon_{t-l_2} + \epsilon_t \Rightarrow \text{Temporally Correlated Process}$$

Side Note: You can essentially do this using autoregressive covariates in a GP

# ARIMA Processes

Differencing in Finance:

$$\text{rev}_t = P_t - P_{t-1}$$

$$\text{net } r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

$$\text{gross } r_t = \frac{P_t}{P_{t-1}}$$

$$\text{risk}_t = \text{Prob}(\text{net } r_t < 0)$$

Differencing Operator:

$$\Delta x_t = x_t - x_{t-1}$$

$$\Delta^2 x_t = \Delta(\Delta x_t) = x_t - 2x_{t-1} + x_{t-2}$$

and so on...

# ARIMA Processes

ARIMA: Integrated ARMA – differences follow an ARMA

$$\mathrm{ARIMA}(p, d, q) : x_t = \Delta^d x_t \sim \mathrm{ARMA}(p, q)$$

# Forecasting

Forecasting = Prediction (everyone has to come up with their own name for things just to be confusing)

$$\text{ARMA}(1,1) : x_{T+1} \mid x_T = \mu + \phi(x_T - \mu) + \theta\epsilon_T + \epsilon_{T+1}$$

$$\epsilon_t \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

$$x_{T+1} \sim \mathcal{N}\left(\mu + \phi(x_T - \mu), \sigma^2(1 + \theta^2)\right)$$

# Forecasting

Forecasting = Prediction (everyone has to come up with their own name for things just to be confusing)

$$\text{ARMA}(1,1) : x_{T+2} \mid x_{T+1} = \mu + \phi(x_{T+1} - \mu) + \theta\epsilon_{T+1} + \epsilon_{T+2}$$

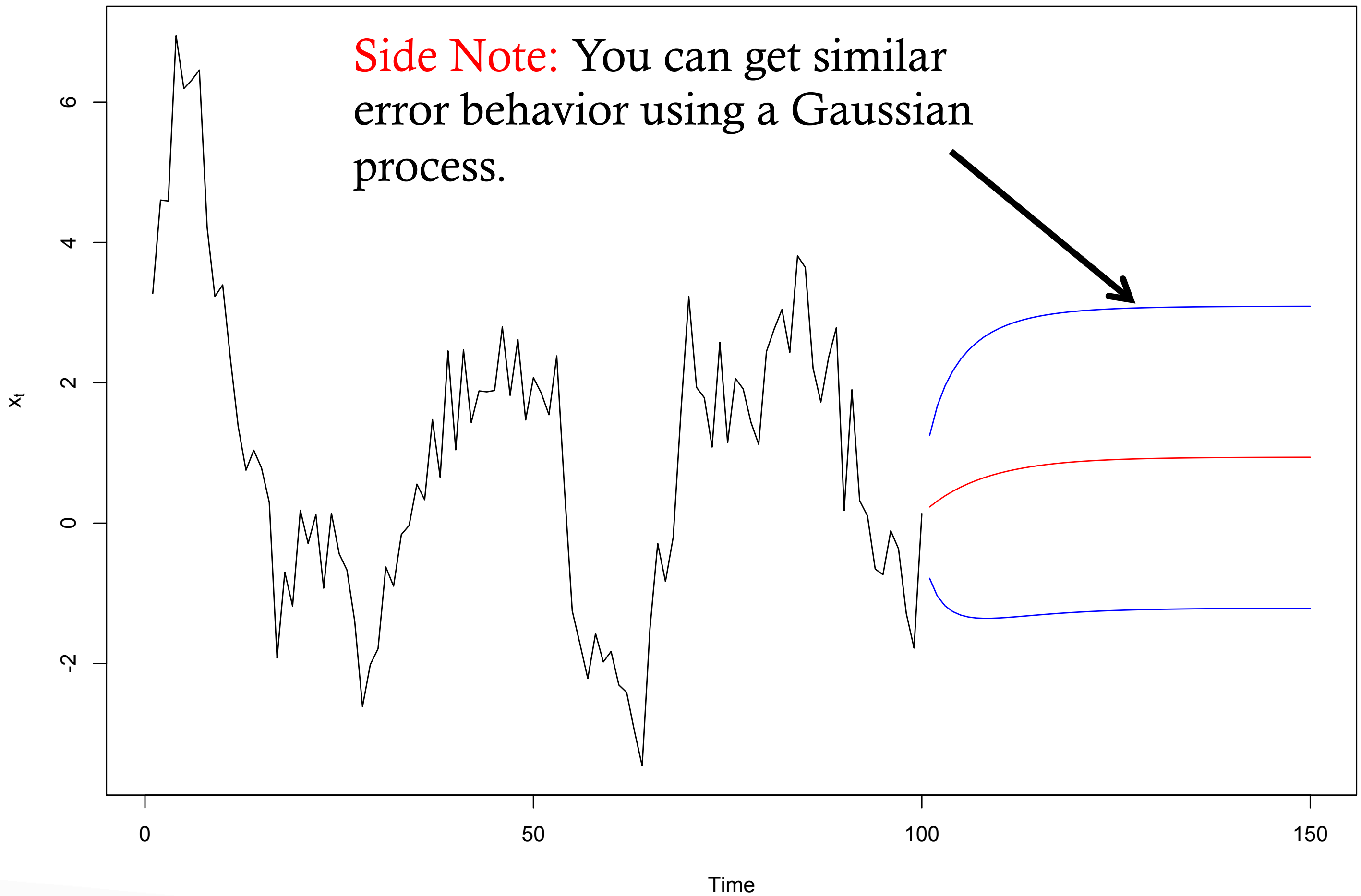
$$\epsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$

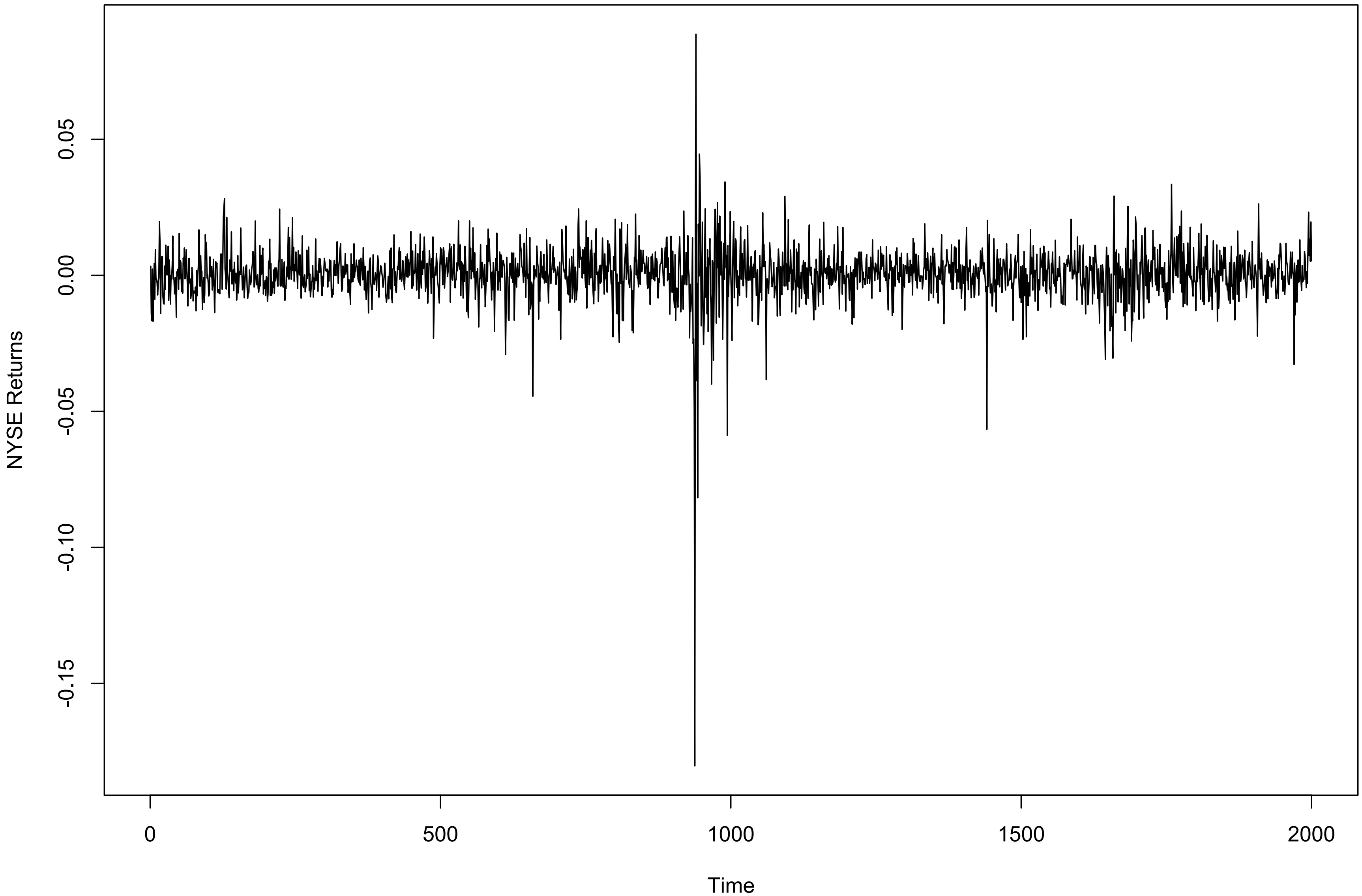
$$x_{T+1} \sim \mathcal{N}(\mu + \phi(x_T - \mu), \sigma^2(1 + \theta^2))$$

$$x_{T+2} \mid x_T \sim \mathcal{N}\left(\mu + \phi^2(x_T - \mu), \sigma^2(1 + \theta^2 + \phi^2\theta^2 + \phi^2)\right)$$

And so on….

# Forecasting






How would you model heteroskedasticity?

# Regression for Std. Deviation

Setup:

$$x_t = \mu_t + \epsilon_t$$
$$\epsilon_t = \sigma_t z_t$$
$$z_t \overset{iid}{\sim} \mathcal{N}(0,1)$$
$$\mathbb{V}\mathrm{ar}(x_t) = \sigma_t^2 \longleftarrow \text{Depends on time}$$

Std. Deviation Regression:

$$\log(\sigma_t) = f(t)$$
$$\text{E.g., } f(t) = \beta_0 + \beta_1 t$$
$$f(t) = \beta_0 + \text{Natural Spline}(t)$$

# Regression for Std. Deviation

Setup:

$$x_t = \mu_t + \epsilon_t$$

$$\epsilon_t = \sigma_t z_t$$

$$z_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$$\mathbb{V}\mathrm{ar}(x_t) = \sigma_t^2 \longleftarrow \text{Depends on time}$$

GP Reg. for Std. Deviation:

$$\log(\sigma_1, \ldots, \sigma_T) \sim \mathrm{GP}(\mu_\sigma, \gamma^2 \rho(\cdot)) \Rightarrow \mathcal{N}(\mu_\sigma \mathbf{1}, \gamma^2 \mathbf{R})$$

$$\log(\hat{\sigma}_1, \ldots, \hat{\sigma}_T) = \max\left[\text{Likelihood} \times \mathcal{N}(\mu_\sigma \mathbf{1}, \gamma^2 \mathbf{R})\right]$$

# ARCH Models

ARCH(1): Autoregressive, conditionally heteroskedastic

$$\epsilon_t = \underbrace{\left(\sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2}\right)}_{\sigma_t} z_t \Rightarrow \mathbb{V}\mathrm{ar}(x_t) = \left(\alpha_0 + \alpha_1 \epsilon_{t-1}^2\right), \alpha_0, \alpha_1 > 0$$

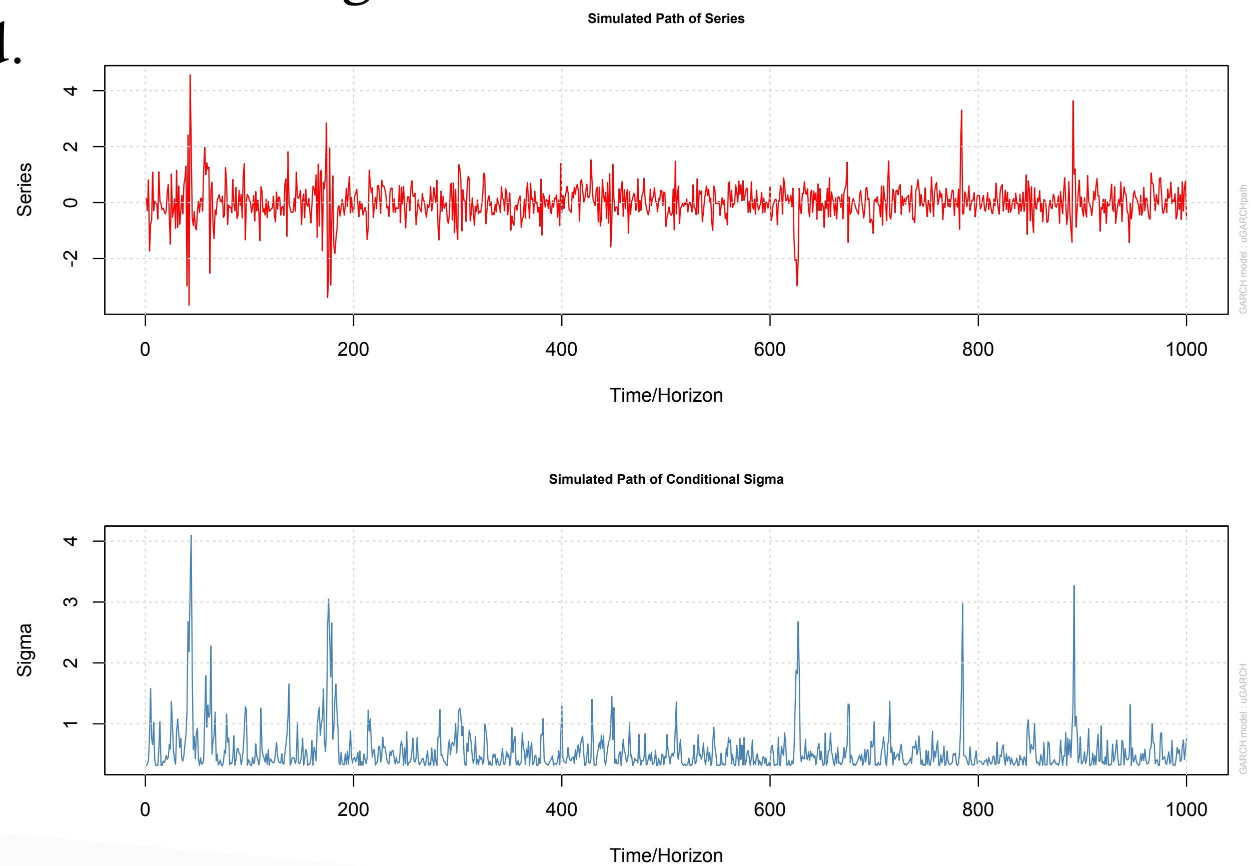
Intuition: If large error at previous time, then variance now increases.

ARCH(q):

$$\epsilon_t = \underbrace{\left(\sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2}\right)}_{\sigma_t} z_t \Rightarrow \mathbb{V}\mathrm{ar}(x_t) = \left(\alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2\right), \alpha_i > 0 \forall i$$

# ARCH Models

Major shortcoming of ARCH:  Inflated variance is short lived.



# GARCH Models

GARCH(p,q):

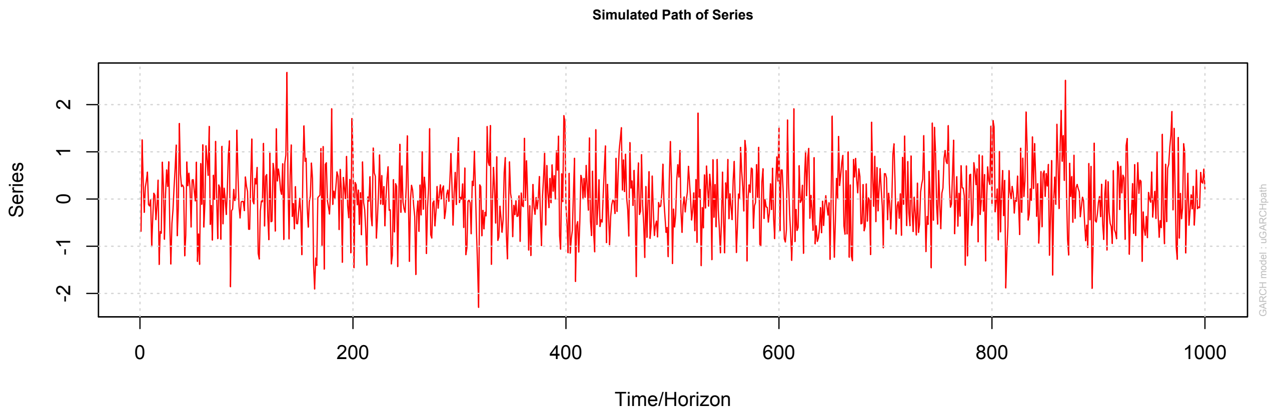
$$\epsilon_t = z_t \sqrt{\alpha_0 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2}$$

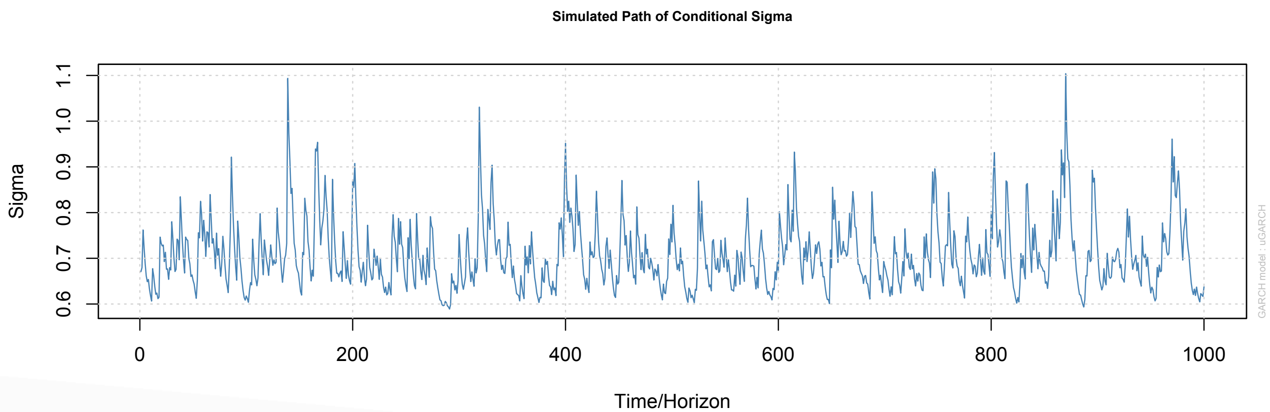
$$\Rightarrow \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$$

$\alpha_0, \alpha_i, \beta_i > 0$

# GARCH Models

GARCH(1,1):



# Regime Switching Models

Big Idea: Time series changes "states" (e.g. growth, recession, stagnation, etc.) across time and behavior is state-specific. Also called a Markov-Switching Model.

$$x_t \mid s_t = k \sim \mathcal{N}\left(\mu_{kt}, \sigma_k^2\right)$$

$$s_t \in \{1, \ldots, K\} : \text{ Latent (hidden) State}$$

$$\mu_{kt} = \text{State Specific Mean}$$

$$\sigma_k^2 = \text{State Specific Variance}$$

$$s_t \sim \text{Markov Chain}(\boldsymbol{\Xi})$$

$$\boldsymbol{\Xi} = \{\xi_{k_1,k_2}\} = \text{Prob}(s_t = k_2 \mid s_{t-1} = k_1)$$

$$\pi_k = \text{Prob}(s_t = k) : \text{ Stationary Probability}$$

$$\Rightarrow x_t \sim \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\mu_{kt}, \sigma_k^2\right) \longleftarrow \text{Mixture Model}$$

# Regime Switching Models

Code for plotting a Gaussian Mixture:

```
## Parameters for a Gaussian Mixture Model
mu <- c(-3,0,3)
s <- c(1,1,1)
prob <- c(1/3,1/3,1/3)

## Function to Calculate Density
mix.dens <- function(x){
as.numeric(dnorm(matrix(x,ncol=length(mu),nrow=length(x)),
mean=matrix(mu,nrow=length(x),ncol=length(mu),byrow=TRUE),
sd=matrix(s,nrow=length(x),ncol=length(mu),byrow=TRUE))%*%prob)
}

## Plot Density
xseq <- seq(mu[1]-3*s[1],mu[3]+3*s[3],length=10000)
ylim <- range(c(dnorm(xseq,mu[1],s[1]),dnorm(xseq,mu[2],s[2]),dnorm(xseq,mu[3],s[3]),mix.dens(xseq)))
plot(xseq,mix.dens(xseq),type="l",lwd=3,xlab="x",ylab='Mixture(x)',ylim=ylim)
lines(xseq,dnorm(xseq,mu[1],s[1]),col="red")
lines(xseq,dnorm(xseq,mu[2],s[2]),col="blue")
lines(xseq,dnorm(xseq,mu[3],s[3]),col="green")
```

# Regime Switching Models

## Code for Drawing from a Gaussian Mixture:

```
## Parameters for a Gaussian Mixture Model
mu <- c(-3,0,3)
s <- c(1,1,1)
prob <- c(1/3,1/3,1/3)

## Draw from a Gaussian Mixture
n <- 10000
z <- sample(1:length(mu),n,replace=TRUE) # Draw which component
draws <- rnorm(n,mean=mu[z],sd=s[z]) # Draw from specific component

## Plot a Histogram of draws with density curve
plot(xseq,mix.dens(xseq),type="l",lwd=3,xlab="x",ylab='Mixture(x)') # see previous page's code
hist(draws,prob=TRUE,add=TRUE)
```

# Regime Switching Models

Inference for Regime Switching Models:

1. Bayesian – Gibbs sampler with no MH steps (easy)
2. Frequentist – Expectation-Maximization (EM) Algorithm (easy if you know the transition probabilities)

The End!