# Lecture 6

## Autocorrelation and Time Series Models

Reading: Forecasting, Time Series, and Regression (4th edition) by Bowerman, O'Connell, and Koehler: Chapter 6

*MATH 4070: Regression and Time-Series Analysis*

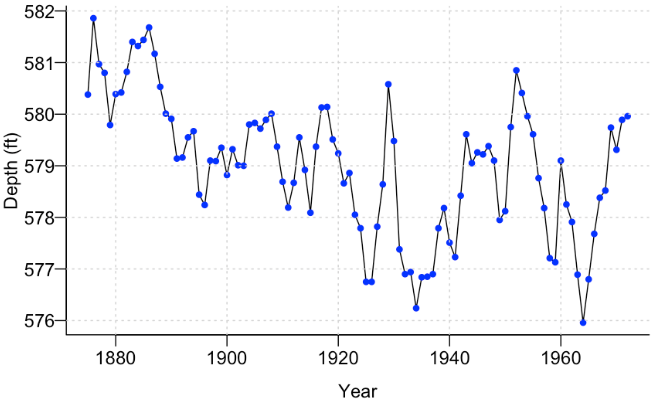
Whitney Huang
Clemson University

# Agenda

1. Objectives of Time Series Analysis
2. Time Series Models
3. Mean and Autocovaraince Functions
4. Stationarity

# Level of Lake Huron 1875–1972

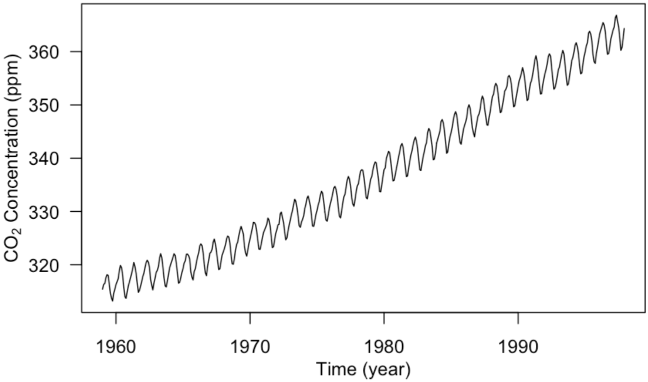
```{r}
par(mar = c(3.2, 3.2, 0.5, 0.5), mgp = c(2, 0.5, 0), bty = "L")
data(LakeHuron)
plot(LakeHuron, ylab = "Depth (ft)", xlab = "Year", las = 1)
points(LakeHuron, cex = 0.8, col = "blue", pch = 16)
grid()
```



# Mauna Loa Atmospheric $CO_2$ Concentration

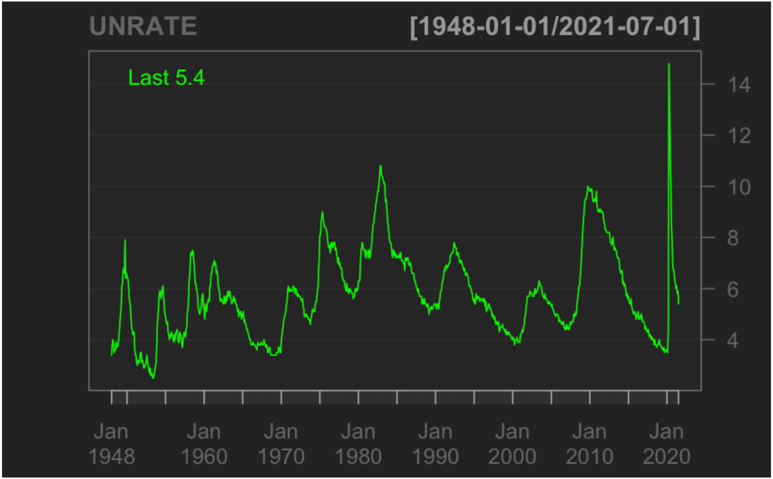


```{r}
data(co2)
par(mar = c(3.8, 4, 0.8, 0.6))
plot(co2, las = 1, xlab = "", ylab = "")
mtext("Time (year)", side = 1, line = 2)
mtext(expression(paste("CO"[2], " Concentration (ppm)")), side = 2, line = 2.5)
```



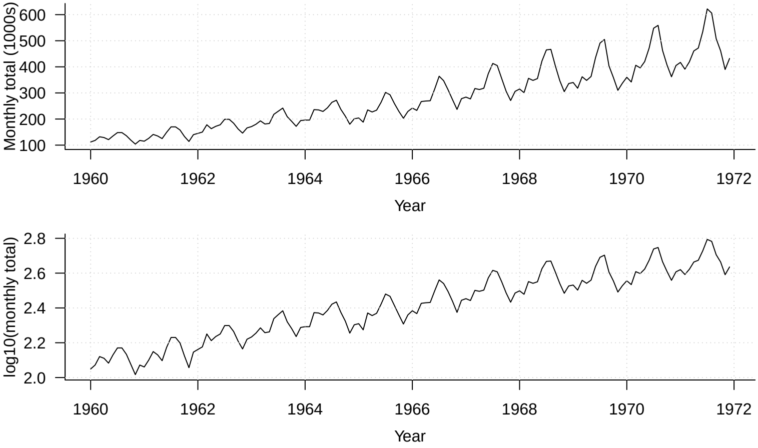
# US Unemployment Rate 1948 Jan. – 2021 July

[Source: St. Louis Federal Reserve Bank's FRED system]



# Airline Passengers Example

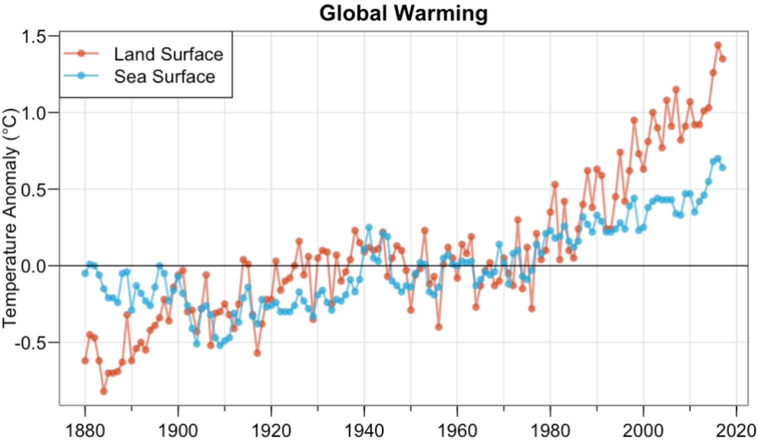
The data set `airpassengers`, which are the monthly totals of international airline passengers from 1960 to 1971.




Here we stabilize the variance with a $\log_{10}$ transformation

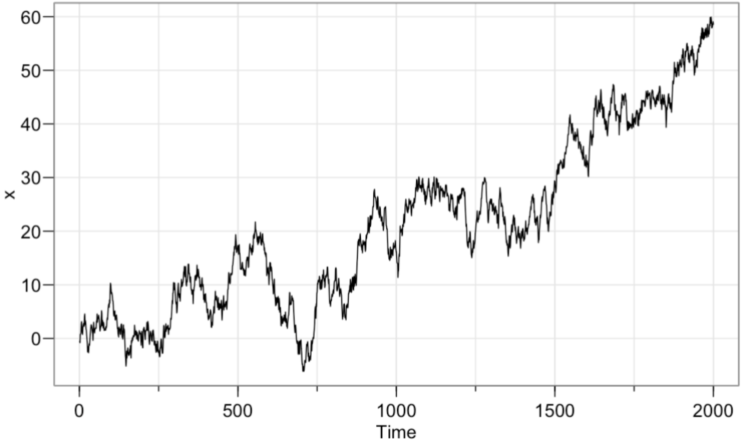
# Global Annual Temperature Anomalies

[Source: NASA GISS Surface Temperature Analysis]



# A Simulated Time Series

```{r}
set.seed(123)
w <- rnorm(2000); x <- cumsum(w); tsplot(x, las = 1)
```



# Objectives of Time Series Analysis

# Some Objectives of Time Series Analysis





**Statistical Modeling**: Find a statistical model that adequately explains the observed time series

- For example, identify a model which can account for the fact that the depths of Lake Huron are correlated with differ years and with a decreasing long-term trend

# Some Objectives of Time Series Analysis



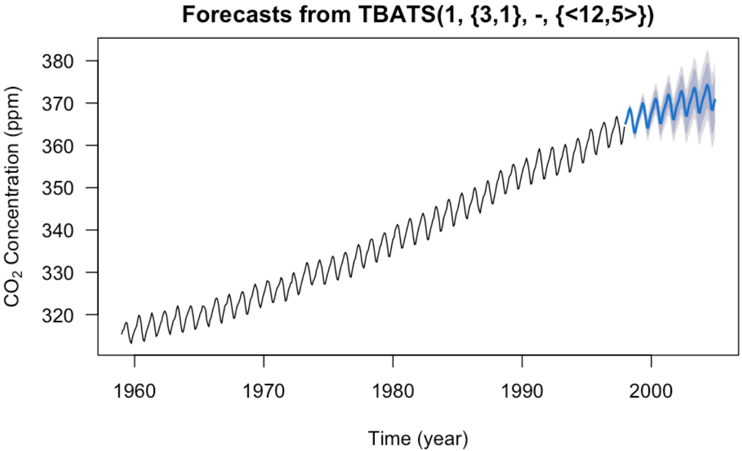


**Statistical Modeling**: Find a statistical model that adequately explains the observed time series

- For example, identify a model which can account for the fact that the depths of Lake Huron are correlated with differ years and with a decreasing long-term trend
- The fitted model can be used for further statistical inference, for instant, to answer the question like: Is there evidence of decreasing trend in the Lake Huron depths?

# Some Objectives of Time Series Analysis, Cont'd

Forecasting is perhaps the most common objective. One observe a time series of given length and wish to **predict** or **forecast** future values of the time series based on those already observed.



# Some Objectives of Time Series Analysis, Cont'd





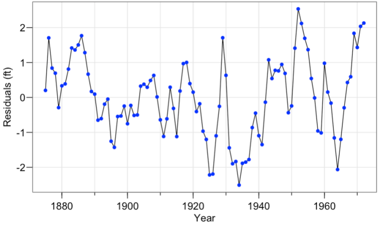


- **Adjustment**: an example would be seasonal adjustment, where the seasonal component is estimated and then removed to better understand the underlying trend
- **Simulation**: use a time series model (which adequately describes a physical process) as a surrogate to *simulate repeatedly in order to approximate how the physical process behaves*
- **Control**: adjust various input (control) parameters to make the time series fit more closely to a given standard (many examples from statistical quality control)

# Time Series Models

# Lake Huron Time Series

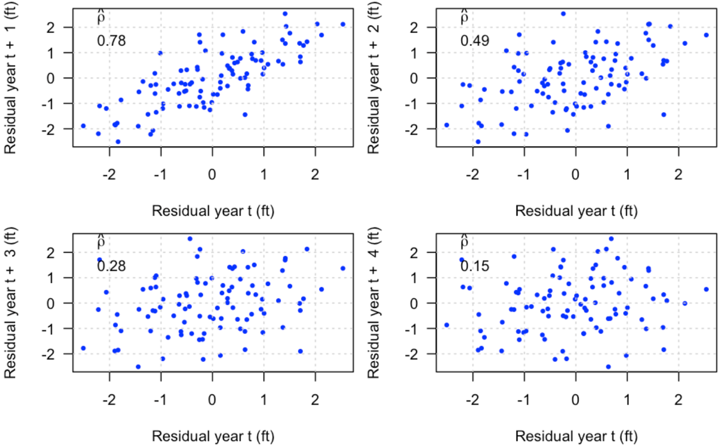
- Time series analysis is the area of statistics which deals with the analysis of dependency between different observations (typically $\{\eta_t\}$)
- Some key features of the Lake Huron time series:
  - decreasing trend
  - some “random” fluctuations around the decreasing trend
- For example, we can extract the ‘noise’ component by assuming a linear trend



# Exploring the Dependence Structure of "Noise" $\{\eta_t\}$

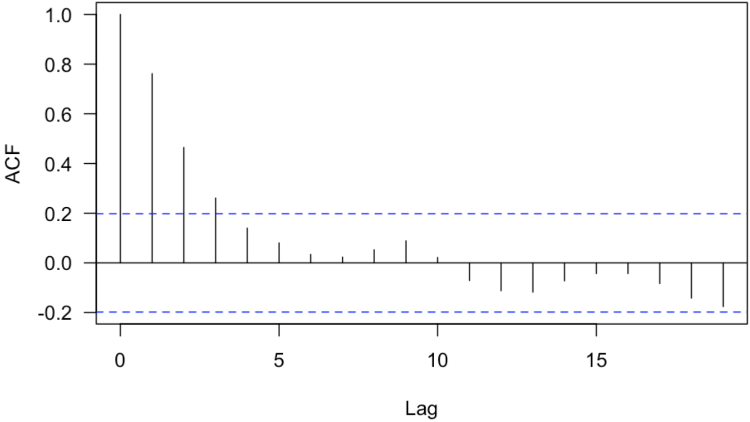


$\{\eta_t\}$ exhibit a temporal dependence structure, meaning that the nearby (in time) values tend to be more alike than those that are far part. To observe this, let's create a few time lag plots



# Further Exploration of the Temporal Dependence Structure

Let’s plot the correlation as a function of the time lag




We will learn how to use this information to suggest an appropriate time series model

# Time Series Models

- A time series model is a probabilistic model that describes how the series data $y_t$ could have been generated. More specifically, it is a probability model for $\{Y_t : t \in T\}$, a collection of random variables indexed in time

## Time Series Models

- A time series model is a probabilistic model that describes how the series data $y_t$ could have been generated. More specifically, it is a probability model for $\{Y_t : t \in T\}$, a collection of random variables indexed in time

- We will keep our models for $Y_t$ as simple as possible by assuming stationarity, meaning that some characteristics of the distribution of $Y_t$ depend only on the "time lag" not on the specific time points

# Time Series Models

- A time series model is a probabilistic model that describes how the series data $y_t$ could have been generated. More specifically, it is a probability model for $\{Y_t : t \in T\}$, a collection of random variables indexed in time
- We will keep our models for $Y_t$ as simple as possible by assuming stationarity, meaning that some characteristics of the distribution of $Y_t$ depend only on the "time lag" not on the specific time points
- While most time series are not stationary, we can model the non-stationary parts (e.g., by **de-trending** or **de-seasonalizing**) to obtain a stationary component, $\eta_t$. We typically assume the process is second-order stationary, meaning

$$\mathbb{E}[\eta_t] = 0, \quad \forall t \in T \quad \text{and,}$$
$$\mathrm{Cov}(\eta_t, \eta_{t'}) = \gamma(t' - t) = \mathrm{Cov}(\eta_{t+s}, \eta_{t'+s})$$

# Time Series Models

- A time series model is a specification of the probabilistic distribution of a sequence of random variables (RVs) $\eta_t$

  (The observed time series is a realization of such a sequence of random variables)

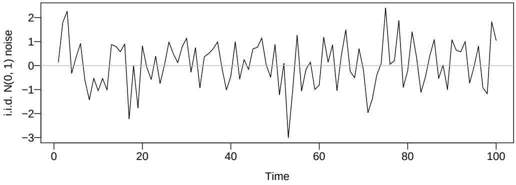
- The simplest time series is i.i.d. (*independent and identically distributed*) noise

  - $\{\eta_t\}$ is a sequence of independent and identically distributed zero-mean (i.e., $\mathbb{E}(\eta_t) = 0, \forall t$) random variables $\Rightarrow$ no temporal dependence
  - It is of little value of using i.i.d. noise model to conduct forecast as there is no information from the past observations
  - **But**, we will use i.i.d. model as a building block to develop time series models that can accommodate time dependence

# Example Realizations of i.i.d. Noise

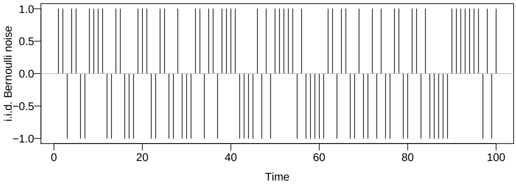
- Gaussian (normal) i.i.d. noise with mean $0$ and variance $\sigma^2 > 0$

$$f(\eta_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{\eta_t^2}{2\sigma^2})$$




- Bernoulli i.i.d. noise with “success” probability

$$\mathbb{P}(\eta_t = 1) = p = 1 - \mathbb{P}(\eta_t = -1)$$



# Means and Autocovarainces

A time series model could also be a specification of the means and autocovariances of the RVs

- The mean function of $\{\eta_t\}$ is

$$\mu_t = \mathbb{E}(\eta_t).$$

- $\mu_t$ is the population mean at time $t$, which can be computed as:

$$\mu_t = \begin{cases} \int_{-\infty}^{\infty} \eta_t f(\eta_t)\, d\eta_t & \text{when } \eta_t \text{ is a continuous RV;} \\ \sum_{-\infty}^{\infty} \eta_t p(\eta_t), & \text{when } \eta_t \text{ is a discrete RV,} \end{cases}$$

where $f(\cdot)$ and $p(\cdot)$ are the probability density function and probability mass function of $\eta_t$, respectively



School of MATHEMATICAL AND STATISTICAL SCIENCES Clemson University

# Examples of Mean Functions

- **Example 1**: What is the mean function for $\{\eta_t\}$, an i.i.d. $\mathrm{N}(0, \sigma^2)$ process?

- **Example 2**: For each time point, let $Y_t = \beta_0 + \beta_1 t + \eta_t$ with $\beta_0$ and $\beta_1$ some constants and $\eta_t$ is defined above. What is $\mu_Y(t)$?

# Review: The Covariance Between Two RVs

- The covariance between the RVs $X$ and $Y$ is

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}\{(X-\mu_X)(Y-\mu_Y)\}\\ &= \mathbb{E}(XY)-\mu_X\mu_Y.\end{aligned}$$

  It is a measure of linear dependence between the two RVs. When $X = Y$ we have

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X).$$

- For constants $a, b, c$, and RVs $X$, $Y$, $Z$:

$$\begin{aligned}\mathrm{Cov}(aX+bY+c,Z) &= \mathrm{Cov}(aX,Z)+\mathrm{Cov}(bY,Z)\\ &= a\mathrm{Cov}(X,Z)+b\mathrm{Cov}(Y,Z)\end{aligned}$$

  $\Rightarrow$

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{Cov}(X,X)+\mathrm{Cov}(X,Y)+\mathrm{Cov}(Y,X)+\mathrm{Cov}(Y,Y)\\ &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)\end{aligned}$$

# Autocovariance Function

- The autocovariance function of $\{\eta_t\}$ is

$$\gamma(s,t) = \mathrm{Cov}(\eta_s, \eta_t) = \mathbb{E}[(\eta_s - \mu_s)(\eta_t - \mu_t)]$$

  It measures the strength of linear dependence between two RVs $\eta_s$ and $\eta_t$

- **Properties**:
  - $\gamma(s,t) = \gamma(t,s)$ for each $s$ and $t$
  - When $s = t$ we have

$$\gamma(t,t) = \mathrm{Cov}(\eta_t, \eta_t) = \mathrm{Cov}(\eta_t) = \sigma_t^2$$

    the value of the variance function at time $t$
  - $\gamma(s,t)$ is a non-negative definite function (will come back to this later)

# Autocorrelation Function

- The autocorrelation function of $\{\eta_t\}$ is

$$\rho(s,t) = \mathrm{Corr}(\eta_s, \eta_t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}$$

  It measures the "scale invariant" linear association between $\eta_s$ and $\eta_t$

- **Properties**:
  - $-1 \le \rho(s,t) \le 1$ for each $s$ and $t$
  - $\rho(s,t) = \rho(t,s)$ for each $s$ and $t$
  - $\rho(t,t) = 1$ for each $t$
  - $\rho(\cdot,\cdot)$ is a non-negative definite function

# Stationarity

- We typically need “replicates” to estimate population quantities. For example, we use

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

to be the estimate of $\mu_X$, the population mean of the **single** RV, $X$

# Stationarity

- We typically need "replicates" to estimate population quantities. For example, we use

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

to be the estimate of $\mu_X$, the population mean of the **single** RV, $X$

- However, in time series analysis, we have $n = 1$ (i.e., no replication) because we only have one realized value at each time point

# Stationarity

- We typically need "replicates" to estimate population quantities. For example, we use

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

to be the estimate of $\mu_X$, the population mean of the **single** RV, $X$

- However, in time series analysis, we have $n = 1$ (i.e., no replication) because we only have one realized value at each time point
- Stationarity means that some characteristic of $\{\eta_t\}$ does not depend on the time point, $t$, only on the "time lag" between time points **so that we can create "replicates"**

> Next, we will discuss strict stationarity and weak stationarity

# Strictly Stationary Processes







- A time series, $\{\eta_t\}$, is strictly stationary if

$$[\eta_1, \eta_2, \cdots \eta_T] \stackrel{d}{=} [\eta_{1+h}, \eta_{2+h}, \cdots \eta_{T+h}],$$

for all integers $h$ and $T \geq 1 \Rightarrow$ the joint distribution are unaffected by time shifts

- Under such the strict stationarity
  - $\{\eta_t\}$ is identically distributed but not (necessarily) independent
  - $\mu_t = \mu$ is independent of time $t$
  - $\gamma(s, t) = \gamma(s + h, t + h)$, for any $s$, $t$, and $h$

# Weakly Stationary Processes

- $\{\eta_t\}$ is weakly stationary if
  - $\mathbb{E}(\eta_t) = \mu_t = \mu$
  - $\text{Cov}(\eta_t, \eta_{t+h}) = \gamma(t, t+h) = \gamma(h)$, finite constant that can depend on $h$ but not on $t$
- Other names for this type of stationarity include second-order, covariance, wide senese. The quantity $h$ is called the lag
- Weak and strict stationarity
  - A strictly stationary process $\{\eta_t\}$ is also weakly stationary as long as $\mu$ is finite
  - Weak stationarity does not imply strict stationarity!

# Autocovariance Function of Stationary Processes

The autocovariance function (ACVF) of a stationary process $\{\eta_t\}$ is defined to be

$$\begin{aligned}\gamma(h) &= \mathrm{Cov}(\eta_t, \eta_{t+h}) \\ &= \mathbb{E}[(\eta_t - \mu)(\eta_{t+h} - \mu)],\end{aligned}$$

which measures the lag-h time dependence

**Properties of the ACVF**:

- $\gamma(0) = \mathrm{Var}(\eta_t)$
- $\gamma(-h) = \gamma(h)$ for each $h$
- $\gamma(s-t)$ as a function of $(s-t)$ is non-negative definite

# Autocorrelation Function of Stationary Processes





The autocorrelation function (ACF) of a stationary process $\{\eta_t\}$ is defined to be

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

which measures the “scale invariant” lag-h time dependence

**Properties of the ACF**:

- $-1 \leq \rho(h) \leq 1$ and $\rho(0) = 1$ for each h
- $\rho(-h) = \rho(h)$ for each $h$
- $\rho(s-t)$ as a function of $(s-t)$ is non-negative definite

# Summary





In this lecture, we discuss

- Objectives of time series analysis
- Time series models
- Mean and auto-covariance/correlation functions
- Stationarity assumption in time series

The most important `R` function of this lecture is `acf`, which calculates and plots the sample autocorrelation