Unit 6 Live Session (Solutions)

Time Series Analysis Lecture 1

Figure 1: South Hall

Class Announcements

- Congratulations on finishing the first part of the course!
- HW 6 is out this week
- Lab-2 is due in 5 weeks

Roadmap

Rearview Mirror

- How to model different types of cross-sectional data
- How to conduct a thorough statistical analysis for binary, unordered multiclass, ordered multiclass, and count data

Today

- Introduction to time series analysis
- Basic terminology of time series analysis
- Notion and measure of dependency
- Notion of stationarity and Ergodic theorem
- Examples of simple time series models

Looking Ahead

- Linear time-trend regression
- Time-series smoothing techniques

Start-up Code

```
### Load a set of packages inclusing: broom, cli, crayon, dbplyr , dplyr, dtplyr, forcats,
## googledrive, googlesheets4, ggplot2, haven, hms, httr, jsonlite, lubridate , magrittr,
## modelr, pillar, purrr, readr, readxl, reprex, rlang, rstudioapi, rvest, stringr, tibble,
## tidyr, xml2
library(tidyverse)

# Insert the function to *tidy up* the code when they are printed out
library(knitr)
opts_chunk$set(tidy.opts=list(width.cutoff=60),tidy=TRUE)

# to generate a random walk plot
library(simts)
```

Introduction

Fundamental concepts in Time-series Analysis

Please define the following concepts:

a)- Stochastic process

- What is a deterministic process?

A stochastic process is a sequence of random variables.

$$\{y_i\}_{i=-\infty}^{\infty} = \{.....y_{-3}, y_{-2}, y_{-1}, y_0, y_1, y_2, y_3.....\})$$

A deterministic process is a process in which no randomness is involved in predicting the next values of the process.

b)- Time series

Time series is a stochastic process where the random variable are ordered based on the time index(t).

$$\{y_t\}_{t=-\infty}^{\infty} = \{.....y_{-3}, y_{-2}, y_{-1}, y_0, y_1, y_2, y_3.....\})$$

c)- Realization of a random process

Realization is a sequence of real numbers assigned to each possible value of y_t over an infinite time. Each realization is one possible outcome of the underlying stochastic process that generated the data.

d)- Sample path

It is a finite subset of a realization

$$\{y_1, y_2, ..., y_T\}$$

Fundamental Properties of Time Series

Expectation and Variance

Suppose we have a time series $Y_1, .., Y_T$

- Expectation of a time series is given by:

$$E(Y_t) = \int_{-\infty}^{\infty} y_t f_{Y_t}(y_t) dy_t$$

- The variance of a times series with a stationary mean is defined as:

$$\sigma^2 = E(Y_t - \mu)^2 = \int_{-\infty}^{\infty} (y_t - \mu)^2 f_{Y_t}(y_t) dy_t$$

a)- Compute the expectation of the following time series at (t-1) and (t).

$$X_t = 0.2 + W_t$$
$$Z_t = 0.5 \cdot t + W_t$$

- Where $\{W\}_{t=-\infty}^{\infty}$ is a white noise process with $E(W_t) = 0$ and $E(W_t^2) = \sigma^2$.

$$E(X_t) = E(0.2 + W_t) = 0.2$$
$$E(X_{t-1}) = E(0.2 + W_{t-1}) = 0.2$$

$$E(Z_t) = E(0.5 \cdot t + W_t) = 0.5 \cdot t$$

$$E(Z_{t-1}) = E(0.5 \cdot t - 1 + W_{t-1}) = 0.5 \cdot t - 1$$

b)- Do the time series have a stationary mean?

Expectation of X_t is constant (0.2) over the time, so it is stationary in mean, but the expectation of Z_t increases over the time, and it is not mean stationary.

c)- Compute the variance of X_t and Z_t at t.

$$Var(X_t) = E(X_t - 0.2)^2 = E(W_t^2) = \sigma^2$$

$$Var(Z_t) = E(Z_t - \beta \cdot t)^2 = E(W_t^2) = \sigma^2$$

Both X_t and Z_t have a constant variance σ^2 over time

The Autocorrelation and Autocovariance Functions

- The autocovariance function of a covariance stationary time series model Y_t is defined as:

$$\gamma_k = Cov(Y_t, Y_{t-k}) = E(Y_t - \mu)(Y_{t-k} - \mu)$$

- The autocorrelation function is a useful measure of how long effects persist in a time series and is defined as:

$$\rho_k = Corr(Y_{t-k}, Y_t) = \frac{\gamma_k}{\gamma_0}$$

a)- What does $\gamma(k) = 0$ imply?

Autocovariance and Autocorrelation are only appropriate to measure a linear dependence, and $\gamma(k) = 0$ does not imply that Y_{t-k} and Y_t are independent but simply that they are uncorrelated.

b)- Compute the γ_1 of X_t and Z_t between t and $t-1$.

$$\gamma_1 = Cov(X_{t-1}, X_t) = E(X_{t-1} - \mu_{t-1})(X_t - \mu_t) = E(X_{t-1} - 0.2)(X_t - 0.2) = E(W_{t-1}W_t) = 0$$

$$\gamma_1 = Cov(Z_{t-1}, Z_t) = E(Z_{t-1} - \mu_{t-1})(Z_t - \mu_t) = E(Z_{t-1} - 0.5 \cdot t - 1)(Z_t - 0.5 \cdot t) = E(W_{t-1}W_t) = 0$$

Both X_t and Z_t $gamma_1$ is not a function of time. So the two consecutive observations are uncorrelated in these two time series. Moreover, if you compute the autocovariance of lag two or lag three etc., all of them will be zero.

c)- Compute the ρ_1 of X_t and Z_t between t and $t-1$.

$$\rho_1 = Corr(X_{t-1}, X_t) = \frac{\gamma_1}{\gamma_0} = \frac{0}{\sigma^2} = 0$$

$$\rho_1 = Corr(Z_{t-1}, Z_t) = \frac{\gamma_1}{\gamma_0} = \frac{0}{\sigma^2} = 0$$

d)- What is the partial autocorrelation function? How is it different from the autocorrelation function?

The Partial Autocorrelation Function can be interpreted as a different version of the ACF which measures autocorrelation between Y_t and Y_{t-k} removing the influence of all observations in between.

For example, let us assume we have three correlated variables Y_t, Y_{t+1} and Y_{t+2} and we wish to find the direct correlation between Y_t and Y_{t+2}. We first apply a linear regression of Y_t on Y_{t+1} (to obtain $\widehat{Y_t}$) and Y_{t+2} on Y_{t+1} (to obtain $\widehat{Y_{t+2}}$) and then we compute

compute $corr(\widehat{Y_t}, \widehat{Y_{t+2}})$ which represents the partial autocorrelation. In this manner, the effect of Y_{t+1} on the correlation between Y_t and Y_{t+2} is removed.

Stationarity

- Recall there are two main types of stationarity that are commonly used:

  1- **Strict Stationarity**: The stochastic process is said to be strictly stationary if for every set of time indices $1 \leq t_1 \leq ..... \leq t_k$, the joint distribution of $Y_{t_1}, Y_{t_2}, ..., Y_{t_k}$ is the same as the joint distribution of $Y_{t_1+h}, Y_{t_1+h}, ..., Y_{t_k+h}$.

  2- **Covariance(weak) Stationarity**: A stochastic process is said to be weakly stationary if it has a constant mean and variance, and its covariance function $\gamma(Y_t, Y_{t+h})$ depends only on h (the time difference) and not on t (time itself).

a)- Are X_t and Z_t stationary? Why or why not?

X_t **has constant mean and variance, and its covariance function does not depend on time because it is always zero for all the lags. so** X_t **is covariance stationary.**

Z_t **has a constant variance, and its covariance function does not depend on the time, but because its expectation is increasing the function of the time, it is non-stationary.**

Ergodic theorem and Estimation

- If Y_t is a stationary and ergodic process with $E(Y_t) = \mu$, then:

$$\bar{y} \equiv \frac{1}{T}\sum_{t=1}^{T} y_t \underset{p}{\rightarrow} \mu$$

a)- What is the intuition behind the Ergodic theorem?

The main problem in time series is that there is only one realization of the underlying process in most cases. For example, we only have one realization for the apple stock prices or U.S annual inflation.

But, if the stochastic process is at least weak stationery, and its distribution remains unchanged, then all observations in a given realization are from the same distribution. If Y is also ergodic, then its sample paths will pass through all parts of the sample space, never getting "stuck" in a sub-region, and the sample average will be consistent for $E(Y_t)$.

b)- What is an example of an ergodic process?

i.i.d. is a strictly stationary and ergodic

c)- Is the Ergodic theorem related to any other important theorem you learned in w203?

The ergodic theorem is a generalization of the weak law of large numbers. In Kolmogorov's LLN, the iid assumption rules out serial dependencies between observations. In ergodic theorem, the serial dependencies are allowed provided that it disappears in the long run

So, if Y_t is a stationary ergodic process, then the sample mean is a consistent estimate of the population mean.

Sample Autocovariance Functions

- A natural estimator of the autocovariance function is given by its sample analog:

$$c_k = \frac{1}{T}\sum_{t=1}^{T-k}(y_t - \bar{y})(y_{t-k} - \bar{y})$$

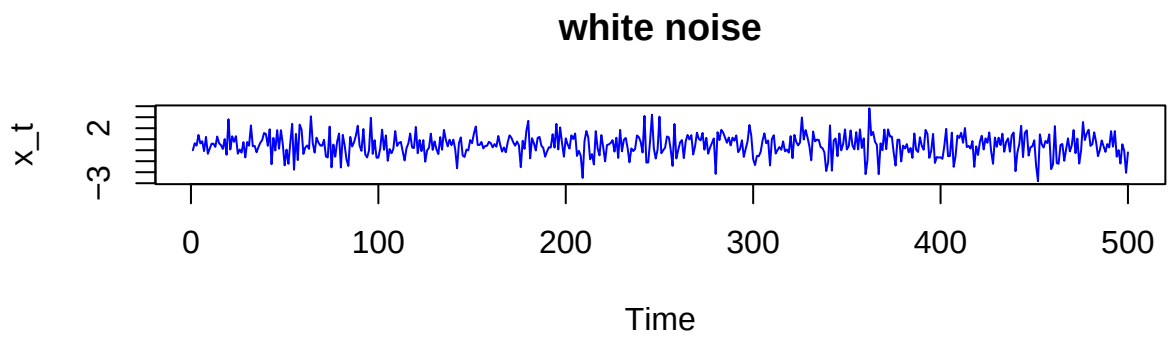
- And the estimator of the autocorrelation function is defined as:

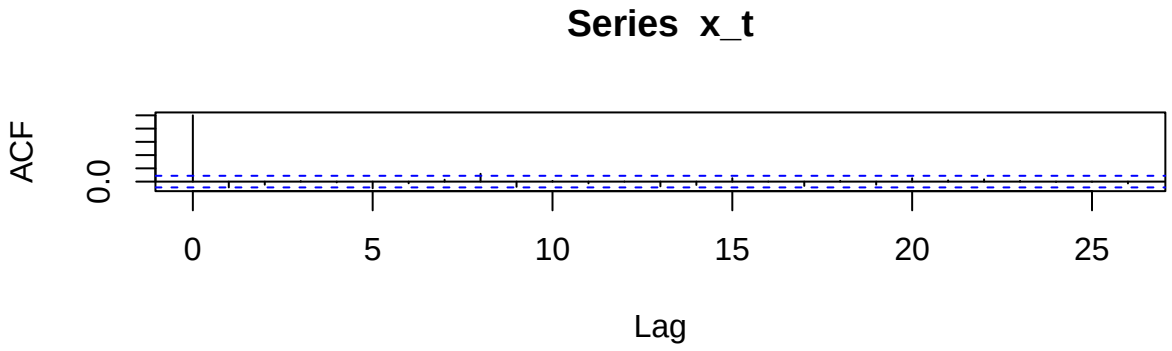
$$r_k = \frac{c_k}{c_0}$$

For a time series, we can plot the sample autocorrelation and sample partial autocorrelation using the ACF and PACF functions in R.

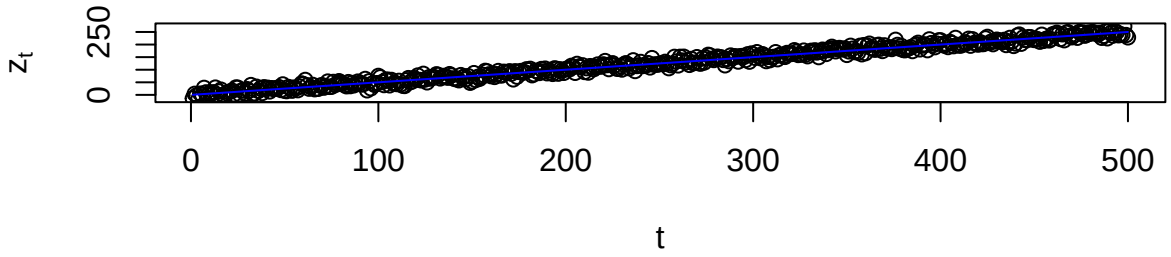
a)- Randomly draw 1000 observations from X_t and Z_t and plot their realizations.

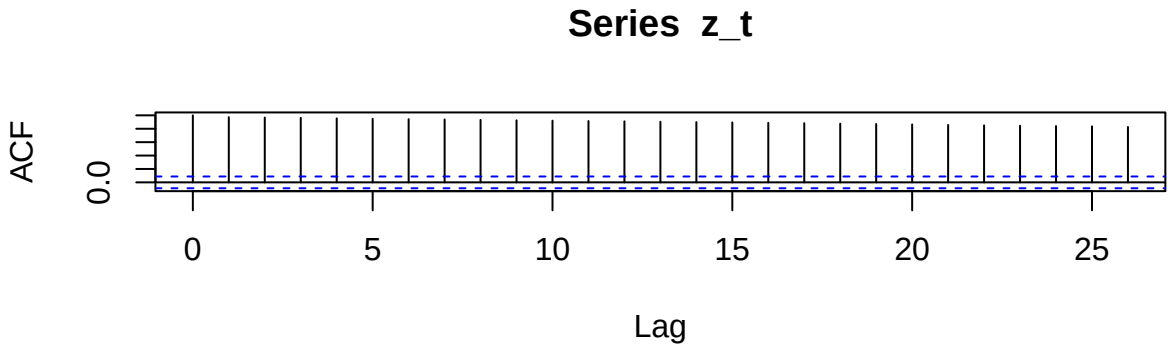
```
set.seed(100)
par(mfrow=c(2,1))

#simulate x_t
alpha = 0.5
x_t = alpha+rnorm(500,0,1)

plot.ts(x_t, main="white noise",col="blue")
acf(x_t)
```






```
##simulate z_t
t=seq(1,500,1)
beta <- 0.5
Tt = beta*t
z_t <-Tt +rnorm(length(t),0,10)

plot.ts(t,z_t,xlab="t",ylab=expression(z[t])); lines(t,Tt,col="blue")
acf(z_t)
```





b)- What do you notice about X_t and Z_t in their realization plots and the ACF?

They are 95% confidence interval of $\rho_k = 0$.

c)- What do you notice about X_t and Z_t? Is there an unexpected pattern in their correlograms?

As we expected, The X_t sample autocorrelations are all zero except at lag eight, which is due to sampling variation.

Because the Z_t is not stationary, the sample mean taken over time is NOT an estimate of the population mean. The population means of Z_t,(0.5t), is a function of time. On the other hand, the sample mean is a scalar, which does not depend on time, and it goes to infinity as the sample size goes to infinity. Computing sample autocorrelation makes sense for a stationary process. However, we can use sample autocorrelations to detect nonstationarity in the data

Stochastic Models

White noise

- A process W_t is white noise if:
  - $E(W_t) = 0$
  - $Var(W_t) = \sigma_2$
  - $cov(W_t, W_{t-k}) = 0$ for $k \neq 0$.
- If W_t is normally distributed, then it is a Gaussian white noise process:
- The white noise model is the building block for most time series models.

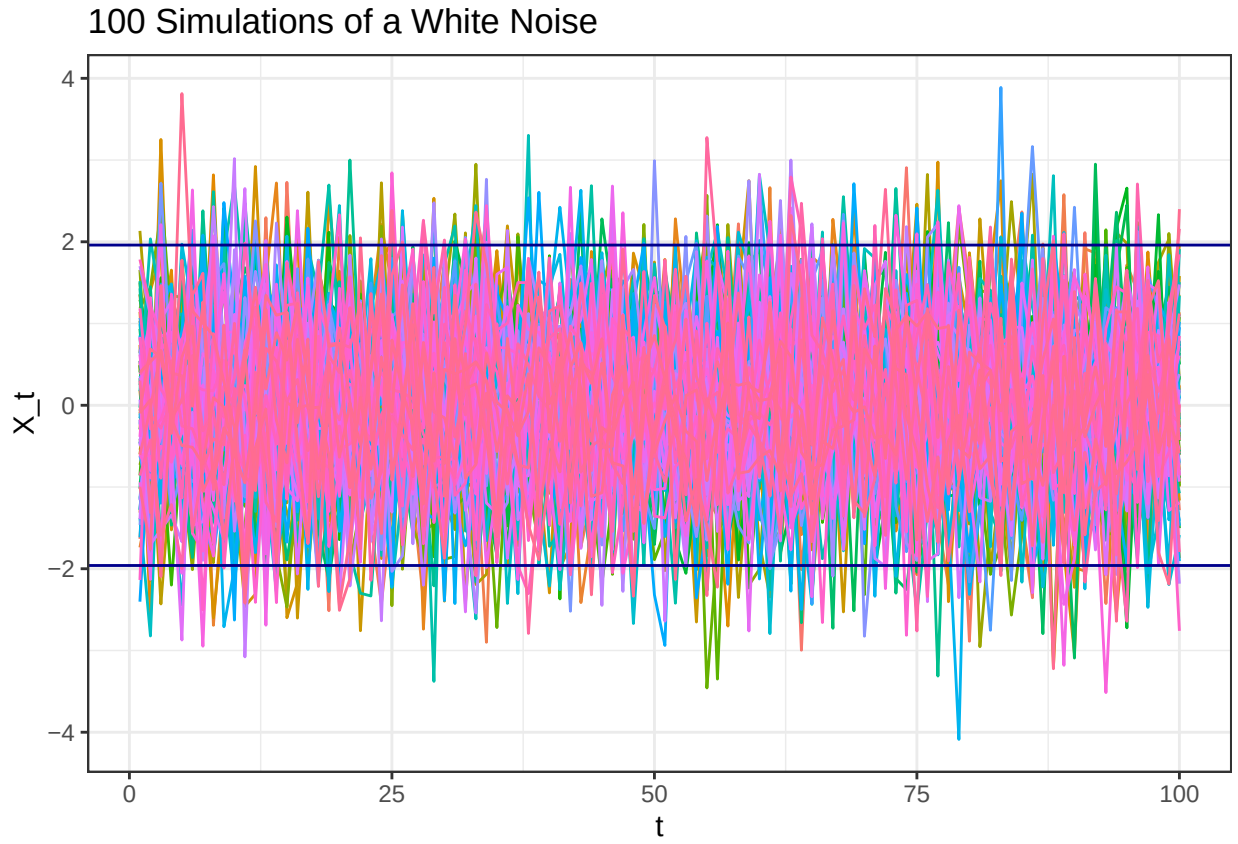
$$W_t \overset{iid}{\sim} N(0, \sigma_w^2)$$

a)- Write a function to take 100 draws from a Gaussian white noise with $\mu = 0$ and $\sigma^2 = 1$. Then use the following code to plot 100 simulations of the white noise process.

```
w <- function(n= 100) {
  w =rnorm(n = n, mean =0, sd = 1)
  return(w)
}

data <- data.frame(t = seq(from = 1, to = 100, by = 1), replicate(w(), n = 100))

data <- data %>%  pivot_longer(
   cols = starts_with("x"),
   names_to = "x",
   values_to = "value"
 )


ggplot(data, aes(x = t, y = value, col = x)) +
   geom_line() +
   ggtitle("100 Simulations of a White Noise") +
   theme_bw() +
   theme(legend.position = "none") +
   xlab("t") + ylab("X_t") +
```

```
  geom_hline(aes(yintercept = -1.96), slope = 0, color = "blue4") +
  geom_hline(aes( yintercept = 1.96), color = "blue4")
```

```
## Warning: Ignoring unknown parameters: slope
```



b)- Do you think this is a stationary time series? Why or why not?

White noise is a stationary process with constant zero mean and constant variance.

Random walk

- A random walk without drift can be defined as:

$$Y_t = Y_{t-1} + W_t$$

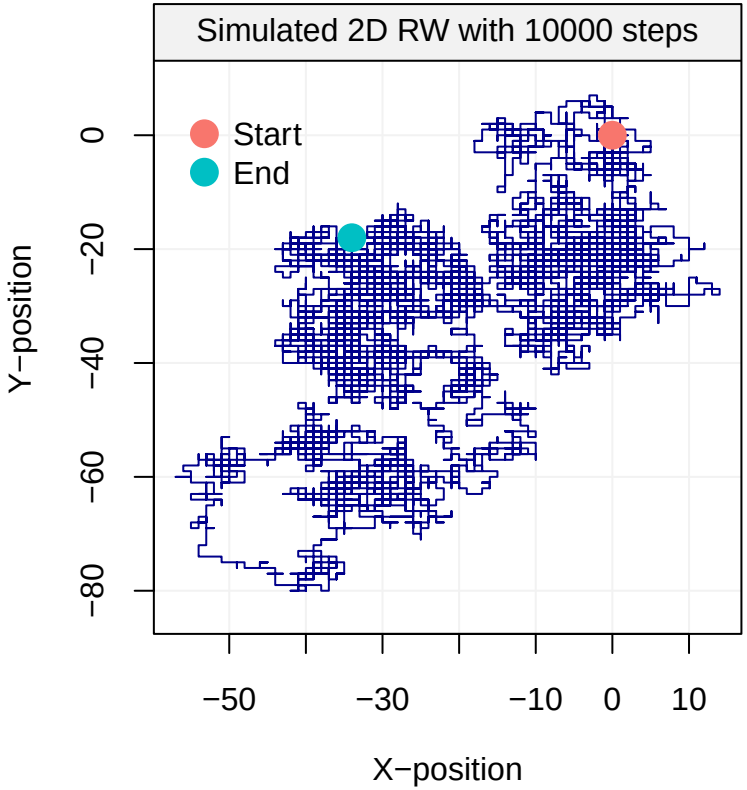
Where W_t is a Gaussian white noise process with initial condition $Y_0 = c$ (usually c=0).

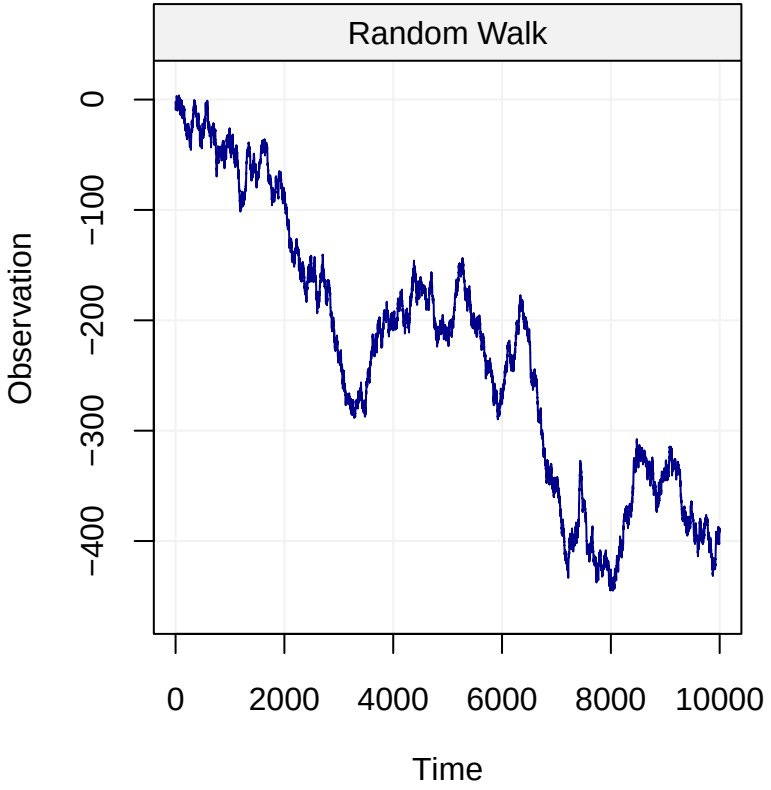
- By back substitution:

$$Y_t = Y_{t-1} + W_t = (Y_{t-2} + W_{t-1}) + W_t = \sum_{i=1}^{t} W_i + Y_0 = \sum_{i=1}^{t} W_i + c$$

- A random walk is the cumulative sum of all the random white noise realizations that preceded it.

- An example of a random walk is a drunk person on Saturday night who is walking on the street, and their next step can either be to their left, right, forward or backward (each with equal probability). The plots below display one realization of such a random walk:





- A random walk with drift is given by:

$$Y_t = \delta + Y_{t-1} + W_t = \delta + (\delta + Y_{t-2} + W_{t-1}) + W_t = \delta \cdot t + \sum_{i=1}^{t} W_i + Y_0 = \delta \cdot t + \sum_{i=1}^{t} W_i + c$$

a)- What is $E(Y_t)$ and $Var(Y_t)$?

$$E(Y_t) = E(Y_{t-1} + W_t) = E(\sum_{i=1}^{t} W_i + c) = \sum_{i=1}^{t} E(W_i) + c = 0 + c = c$$

$$Var(Y_t) = Var(Y_{t-1} + W_t) = Var(\sum_{i=1}^{t} W_i + c) = \sum_{i=1}^{t} Var(W_i) + Var(c) = \sum_{i=1}^{t} \sigma^2 + 0 = t\sigma^2$$

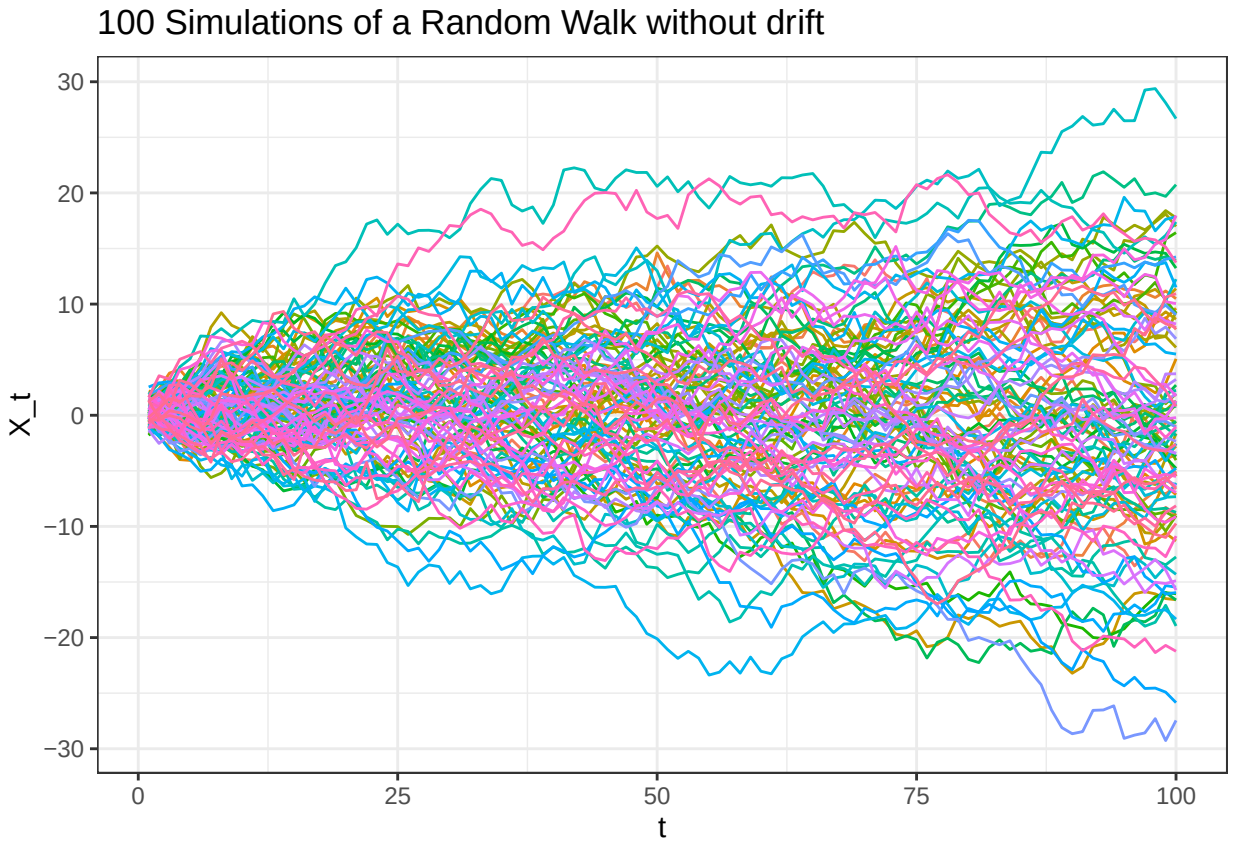
b)- Is a random walk with drift covariance stationary?

Variance of a random walk without drift is increasing function of time, so it is not stationary

c)- Write a function to simulate random walks without drift and $Y_0 = 0$, and plot 100 simulated random walks without drift using the following code.

```
rw_no_drift <- function(n = 1000) {
  w=rnorm(100,0,1)
  xd = cumsum(w)
  return(xd)
}

### After you write the function uncomment and run the code to produce the plot
data <- data.frame(t = seq(from = 1, to = 100, by = 1), replicate(rw_no_drift(), n = 100))

data <- data %>%  pivot_longer(
    cols = starts_with("x"),
    names_to = "x",
    values_to = "value"
  )


ggplot(data, aes(x = t, y = value, col = x)) +
    geom_line() +
    ggtitle("100 Simulations of a Random Walk without drift") +
```

```
    theme_bw() +
    theme(legend.position = "none") +
    xlab("t") + ylab("X_t")
```



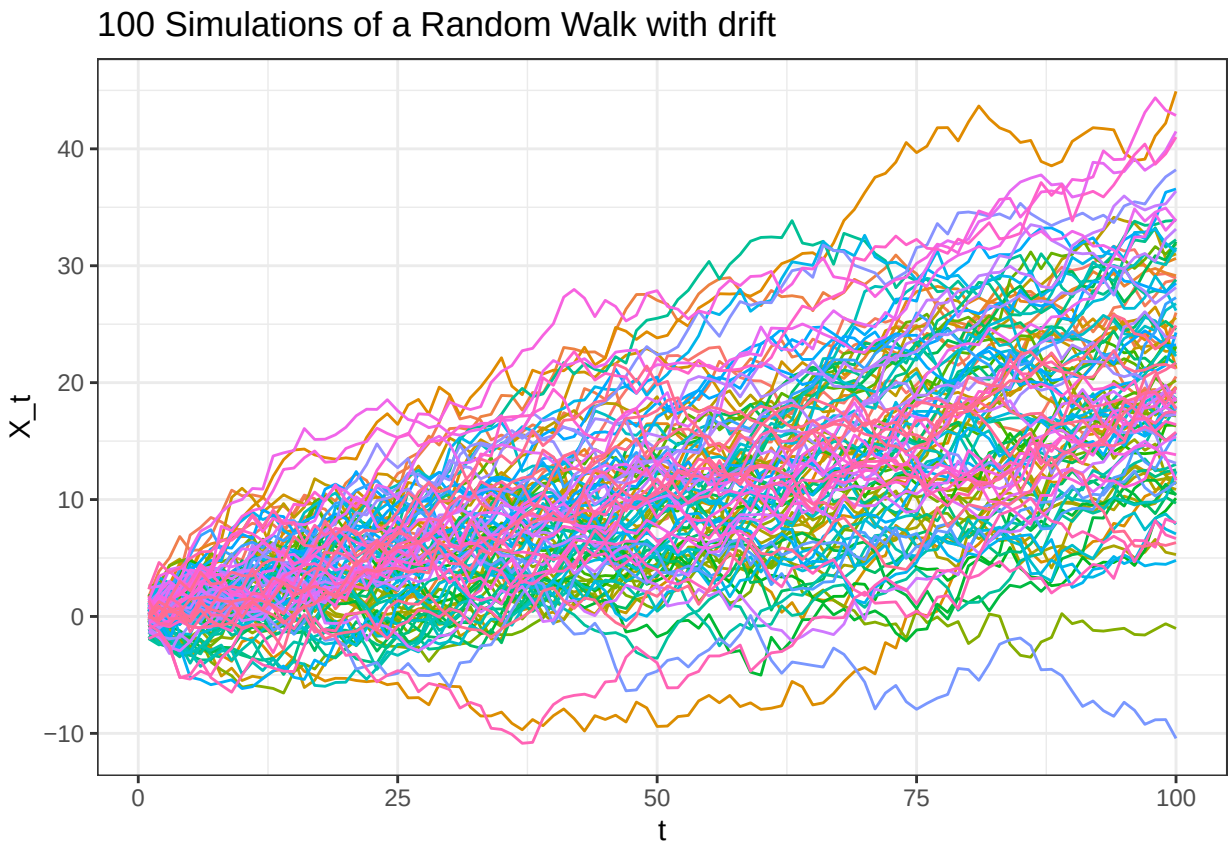
d)- Repeat steps a-c for a random walk with drift $\delta = 0.2$.

$$E(Y_t) = E(\delta + Y_{t-1} + W_t) = E(\delta \cdot t + \sum_{i=1}^{t} W_i + c) = E(\delta \cdot t) + \sum_{i=1}^{t} E(W_i) + c = \delta \cdot t + 0 + c = \delta t \cdot c$$

$$Var(Y_t) = Var(\delta + Y_{t-1} + W_t) = Var(\delta \cdot t + \sum_{i=1}^{t} W_i + c) = Var(\delta \cdot t) + \sum_{i=1}^{t} Var(W_i) + Var(c) = 0 + \sum_{i=1}^{t} \sigma^2 + 0 = t\sigma^2$$

Both expectation and variance of a random walk with drift is increasing function of time, so it is not stationary

```
rw_drift <- function(n = 100) {
  w=rnorm(100,0,1)
  wd = 0.2 + w
  xd = cumsum(wd)
  return(xd) # Replace with your code
}

###After you write the function uncomment and run the code to produce the plot
data <- data.frame(t = seq(from = 1, to = 100, by = 1), replicate(rw_drift(), n = 100))
data <- data %>%  pivot_longer(
     cols = starts_with("x"),
     names_to = "x",
     values_to = "value"
   )


ggplot(data, aes(x = t, y = value, col = x)) +
     geom_line() +
     ggtitle("100 Simulations of a Random Walk with drift") +
     theme_bw() +
     theme(legend.position = "none") +
     xlab("t") + ylab("X_t")
```



e)- What is the main difference between a random walk with and without drift?

While the expectation of a random walk without drift is a function of time, a random walk with drift has both time-variant expectation and variance

First-Order Autoregressive Model

- A first-order autoregressive model or AR(1) is defined as:

$$Y_t = \phi Y_{t-1} + W_t$$

- by back substitution:

$$Y_t = \phi Y_{t-1} + W_t = \phi(\phi_t Y_{t-2} + W_{t-1}) + W_t = \phi^2 Y_{t-2} + \phi W_{t-1} + W_t = \phi^t \cdot Y_0 + \sum_{i=0}^{t-1} \phi^i W_{t-i}$$

- If $|\phi| < 1$, then $lim_{i\to\infty}\phi^i Y_{t-i} = 0$ and the $AR(1)$ process (Y_0=0) is:

$$Y_t = \sum_{i=0}^{t-1} \phi^i W_{t-i}$$

a)- How is an AR(1) process related to white noise and a random walk?

A first-order autoregressive model or AR(1) is a generalization of both the white noise and the random walk processes. So if ϕ is zero, it is white noise, and when ϕ is one, it is a random walk process

b)- Is this AR(1) process stationary?

AR(1) is stationary, if $|\phi| < 1$

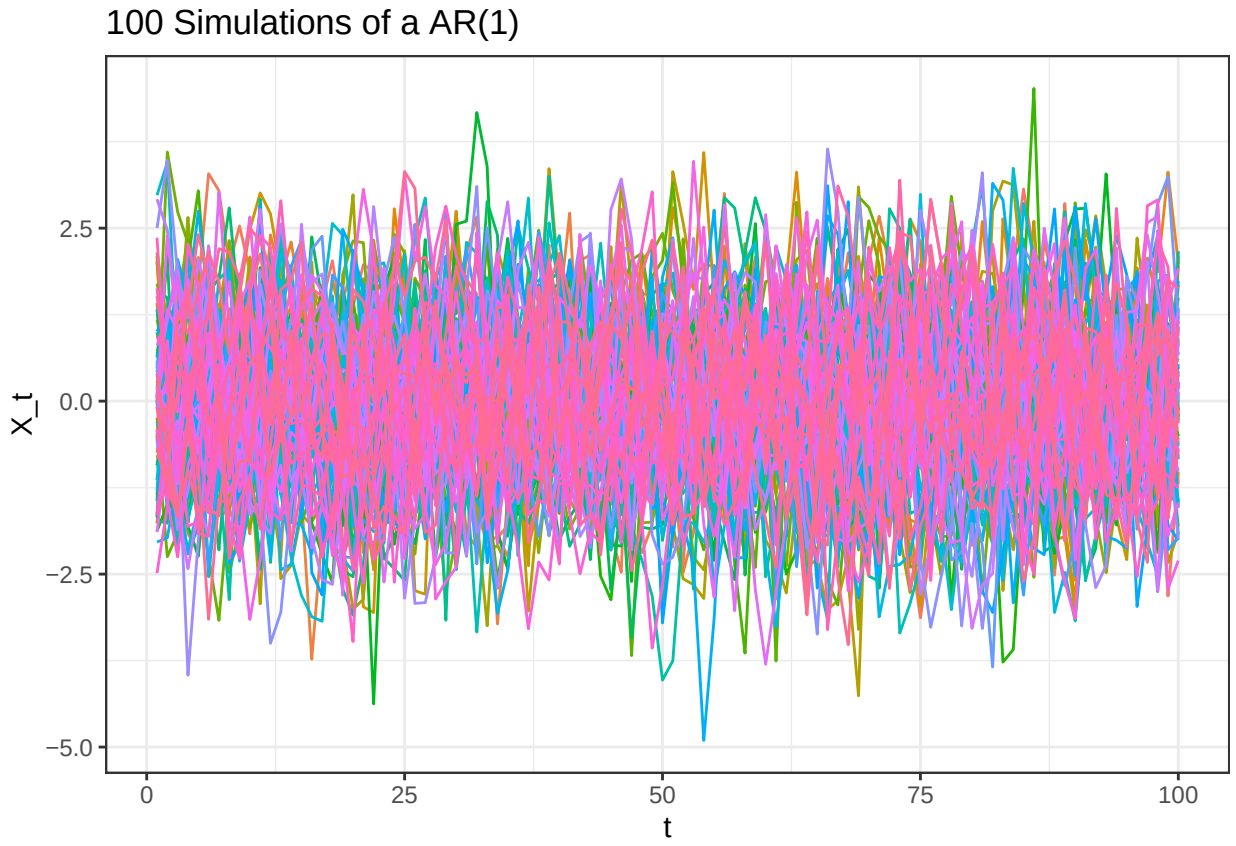
c)- What happens if $\phi = 1$ or $\phi > 1$?

When $\phi = 1$ it is a random walk which is obviously non-stationary, and when $\phi > 1$, it shows exponential growth over the time and is not stationary again.

d)- Use the following code to simulate 100 realizations of an AR(1) process with $\phi = 0.5$.

```
## ar_1()
phi = 0.5

ar_1 <- function(n = 100) {
  y <- w <- rnorm(n, mean = 0, sd = 1)
  for (t in 2:100) y[t] <- phi* y[t - 1] + w[t]
    return(y)
  }

data <- data.frame(t = seq(from = 1, to = 100, by = 1), replicate(ar_1(), n = 100))
```

```r
data <- data %>%  pivot_longer(cols = starts_with("x"),
                               names_to = "x",values_to = "value")
ggplot(data, aes(x = t, y = value, col = x)) +
     geom_line() +
     ggtitle("100 Simulations of a AR(1)") +
     theme_bw() +
     theme(legend.position = "none") +
     xlab("t") + ylab("X_t")
```



Moving Average Process of Order 1

- An AR(1) can be written as a linear combination of all past white noise (Wt). This is known as the invertibility of AR(1) processes.
- Similarly, an MA(1) can be written as a linear combination of the white noises, but it is a "truncated" version and only includes two white noise terms.

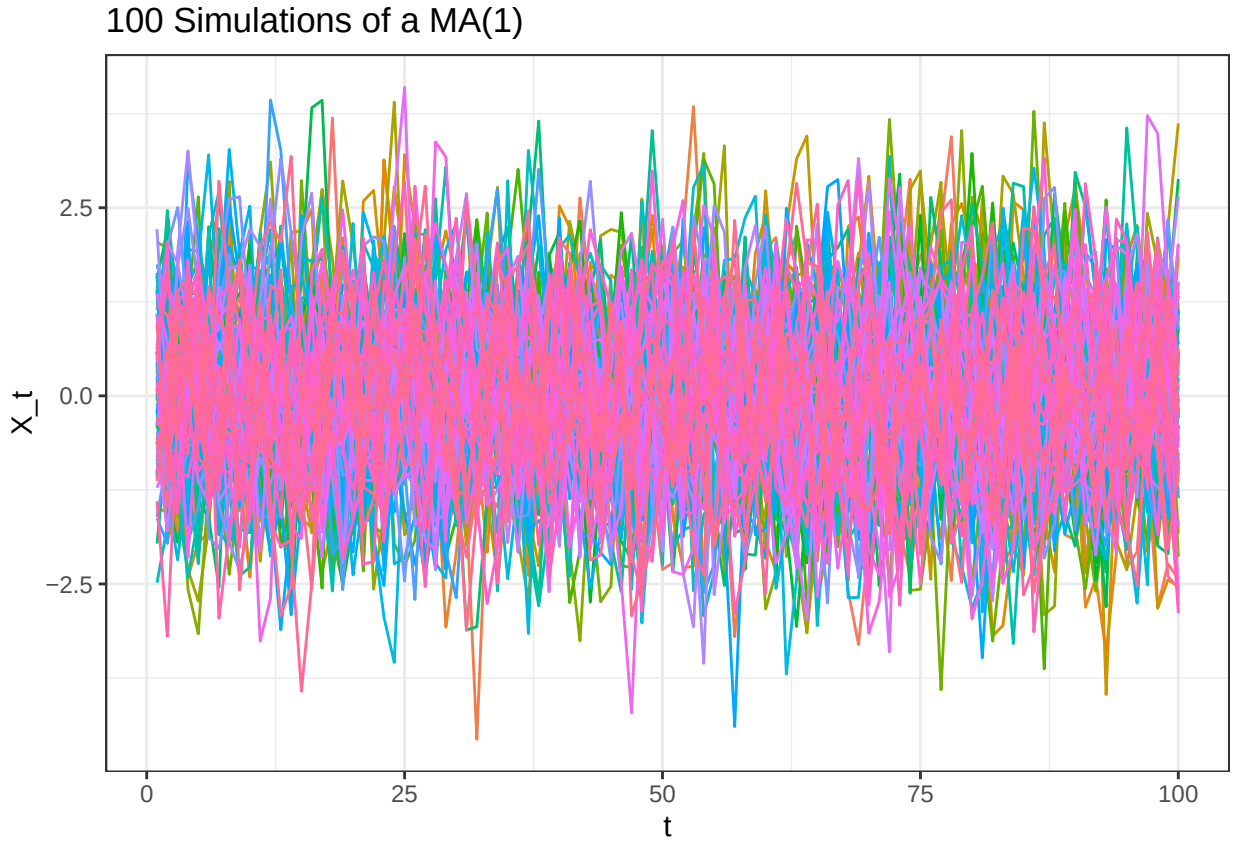
$$Y_t = \theta W_{t-1} + W_t$$

a)- Is the MA(1) process stationary?

MA(1) has a constant expectation and variance over the time, so its a stationary process

b)- Run following code to simulate 100 realizations of a MA(1) model with $\phi = 0.5$.

```
## ma_1()
theta = 0.5
ma_1 <- function(n = 100) {
      y <- w <- rnorm(n, mean = 0, sd = 1)
  for (t in 2:100) y[t] <- theta* w[t - 1] + w[t]
  return(y)
}

data <- data.frame(t = seq(from = 1, to = 100, by = 1), replicate(ma_1(), n = 100))

data <- data %>%  pivot_longer(
     cols = starts_with("x"),
     names_to = "x",
     values_to = "value"
   )


ggplot(data, aes(x = t, y = value, col = x)) +
     geom_line() +
     ggtitle("100 Simulations of a MA(1)") +
     theme_bw() +
     theme(legend.position = "none") +
     xlab("t") + ylab("X_t")
```


100 Simulations of a MA(1)
X_t
2.5
0.0
−2.5
0
25
50
75
100
t

Time Series classes in R

R has several libraries and classes for dealing with time-series data. Of course, it is possible to represent time series in a 'normal' data frame with one column for the time and another for the observations in the series. However, the data will usually be better suited for analysis when it takes the form of one of R's specialized time series classes, which have different properties from regular data frames.

ts objects

- The `ts` object is the most basic type of time-series object in R, requiring only the base `stats` package, which is automatically loaded when you start R.
- `ts` objects come with `frequency`, `start`, and `end` arguments. The `frequency` attribute specifies the number of (regularly-spaced) intervals per unit of time; a frequency of `7` might correspond to daily intervals of weekly units, `52` might correspond to weekly intervals of annual units, while `1` might correspond to annual data or any data where there is only one interval per time-unit.
- The `start` and `end` attributes consist of either a single number specifying a time unit or a vector of two numbers specifying both a time unit and a particular number of intervals in that unit.
- To extract a subset of the time series, you can use the `window()` function with `start` and/or `end` arguments.
- The `time()` function returns the numeric time stamps for the observations.

xts and zoo objects

- `xts` stands for eXtensible Time Series. It is essentially matrix + (time-based) index (aka, observation + time), which allows irregular time intervals.
- `xts` is a constructor or a subclass that inherits behavior from the parent `zoo` (Z's Ordered Observations). It extends the popular `zoo` class, and most `zoo` methods work for xts. These include methods for subsetting, merging, and interpolating time series data.
- `xts` are indexed by a formal time object. Therefore, the data is time-stamped. The two most important arguments are `x` for the data and `order.by` for the index. `x` must be a vector or matrix. `order.by` is a vector of the same length or number of rows of `x`; it must be a proper time or date object and be in increasing order. The `coredata()` and `index()` functions retrieve the observations and their time stamps, respectively.

tsibble objects

- `tsibble` objects are variants of `tibble` objects, which are variants of data frames, and can be used with `dplyr` data-wrangling functions. They require the specification of an index, which can be regular or irregular; in the case of regular intervals, the `yearquarter`, `yearmonth`, `yearweek`, `Date`, and `POSIXct` functions can convert time information to the appropriate class. Tsibble definitions can also include the specification of a critical variable allowing multiple time series to be manipulated as a single object. The `tsibble` library has various functions for subsetting, merging, and interpolating time series data.

Reminders

1. Welcome to the Time Series part of the course!
2. Before the next live session:
    1. Complete the HW-6
    2. Complete all videos and reading for unit 7