1.2 - Time Series Statistical Models

Time Series:

> "A collection of random variables indexed according to the order they are obtained in time."
> Ex: The sequence of random variables $x_1, x_2, x_3, \dots$ where x_t denotes the value taken by the series at time t.

Stochastic Process:

> "A collection of random variables, $\{x_t\}$, indexed by t."
> "The observed values of a stochastic process are referred to as a **realization** of the stochastic process."

Graphing a Time Series:

> The y-axis is typically for the values of the random variables and the x-axis is for the time.
> It is usually convenient to connect the values at adjacent points in time which treats the data as if it is a **continuous time series**.

It is important for time series theory that a continuous parameter time series can be "specified in terms of finite-dimensional distribution functions defined over a finite number of points in time."

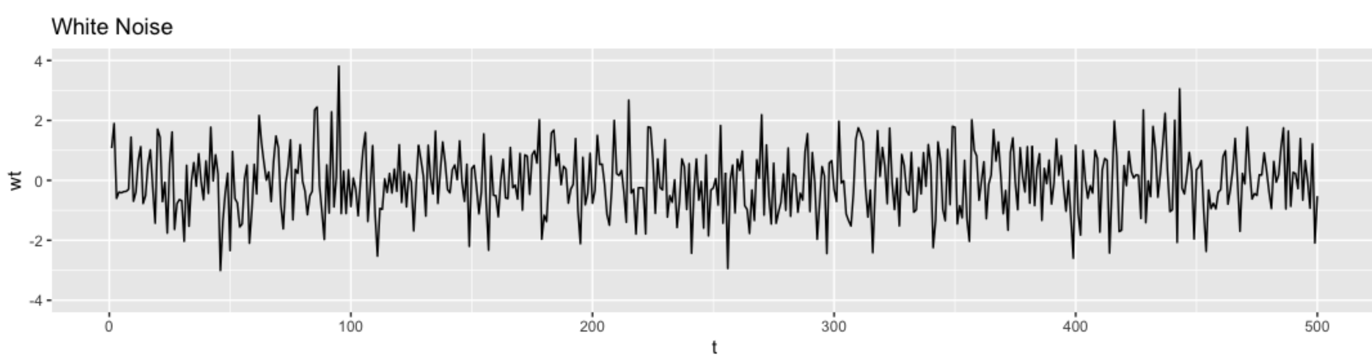
Below we will start seeing examples motivating the use of various combinations of random variables emulating real time series data.

Example 1.8 White Noise

> A collection of uncorrelated random variables, w_t, with mean 0 and finite variance σ_w^2 which is used as a model of noise in engineering applications where it is called *white noise*, and for the purposes of this text we will denote this process $w_t \sim \text{wn}(0, \sigma_w^2)$.
>
> When the noise consists of independent and identically distributed (iid) random variables then we denote this as $w_t \sim \text{iid}(0, \sigma_w^2)$ or "white independent noise" or "iid noise".
>
> When the noise consists of independent normal random variables then we denote this as $w_t \sim$ iid N$(0, \sigma_w^2)$ where $\sigma_w^2 = 1$ or "Gaussian white noise".

```
library(ggplot2)
library(ggfortify)
set.seed(1)

w=rnorm(500)
autoplot(as.ts(w), main="White Noise", xlab="t", ylab="wt", ylim=c(-4,4))
```



Example 1.9 Moving Averages and Filtering

Moving Average

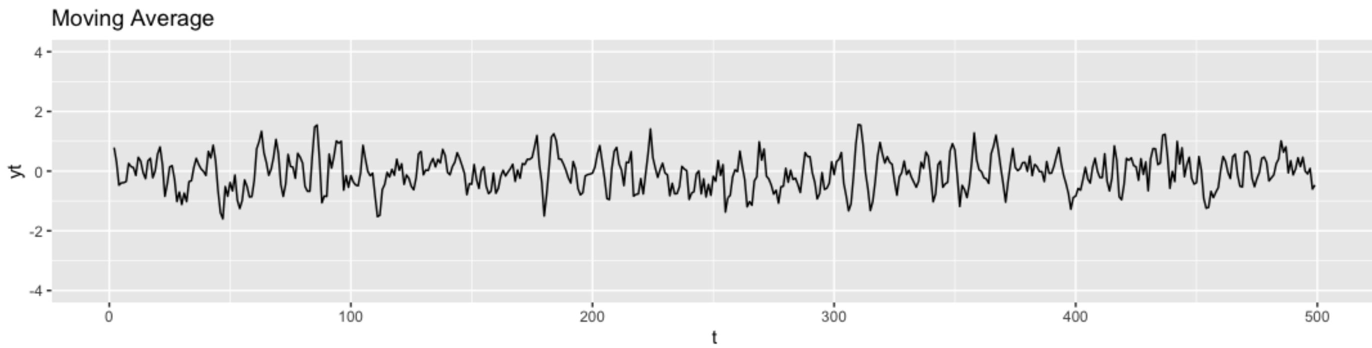
> We replace w_t with v_t where v_t is the average of some fixed window of points around w_t for all w_t in the time series.
> Example: If we fix our window size to be 3, then $v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$ and this means for each point w_t we are replacing it with

> the average of its value and the value of the points before and after it.
> Intuitively, this should mean when we make our window bigger, we should get more smoother time series curve.

Filtering

> Any linear combination of values in a time series. (Moving average like $\frac{1}{3}(w_{t-1}+w_t+w_{t+1})$ is an example)

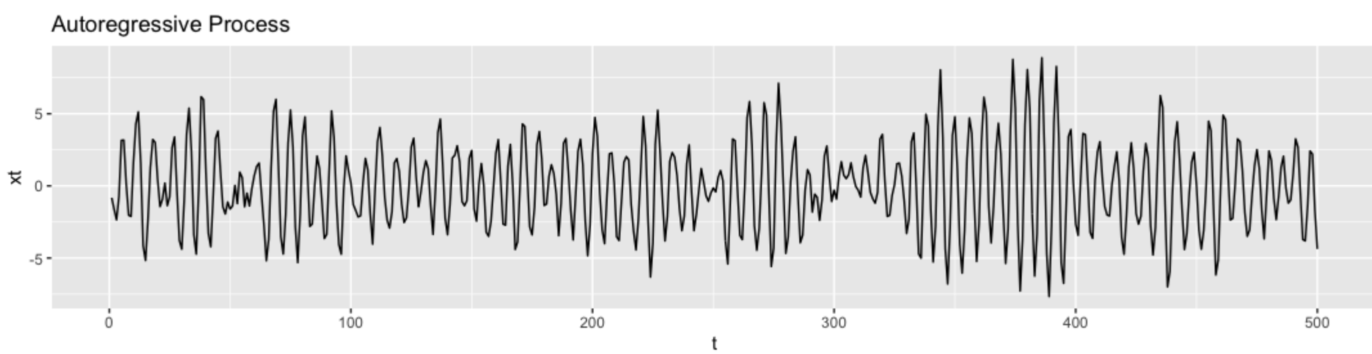
```
v=filter(w, sides=2, filter=rep(1/3,3))
autoplot(as.ts(v), main="Moving Average", xlab='t', ylab="yt", ylim=c(-4,4))
```



Autoregression

> A time series model where the prediction of the current value x_t is a function of some previous values x_{t-k}.
> Ex: $x_t = x_{t-1} + 0.9x_{t-2} + w_t$. Here x_t depends on previous values and another time series w_t.
> Note: We need to have initial values for the time series. (like x_0 and x_{-1} for the example above)

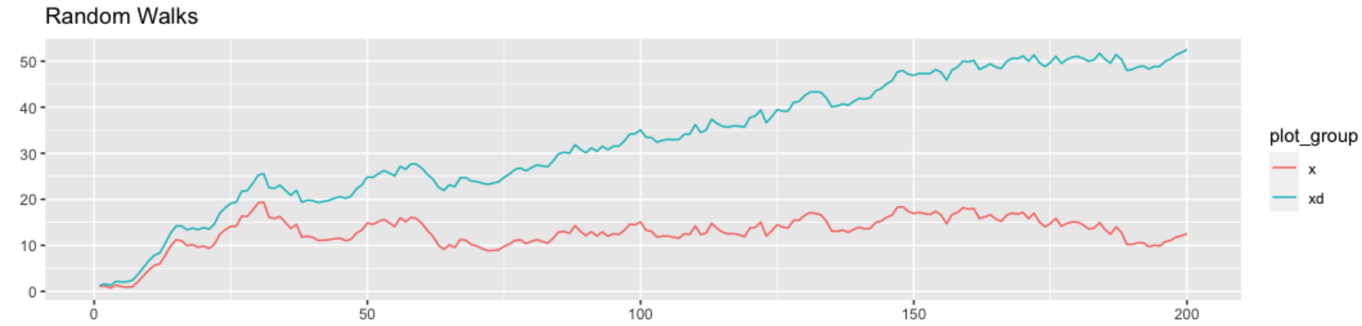
```
w=rnorm(510)
x=filter(w, filter=c(1,-.9), method = "recursive")[-(1:10)]
autoplot(as.ts(x), main="Autoregressive Process", xlab="t", ylab="yt")
```



Random Walk with Drift

> Model of the form $x_t = \delta + x_{t-1} + w_t$ where δ is a drift variable and w_t is white noise.
> We can also write it in the form $x_t = \delta t + \sum_{j=1}^{t} w_j$.
> When $\delta = 0$, we call it just a **random walk**.

```
w=rnorm(200);d=0.2
x=cumsum(w)
wd=w+d
xd=cumsum(wd)
comparison=cbind(x,xd)
autoplot(as.ts(comparison),facets=F,main="Random Walks")
```



Signals with Noise

> Many realistic time series models have an underlying signal with some periodic variation.
> A sinusoidal signal with noise can be modeled in the form $x_t = A\cos(2\pi\omega t + \phi) + w_t$ where A is the amplitude, ω is the frequency of the oscillation, ϕ is a phase shift, and w_t is white noise.
> Ratio of amplitude of the signal to σ_w is called the **signal to noise ratio** (SNR).
> The larger the SNR, the easier it is to detect the signal.