Time Series

-Quantifying the World-

Lecturer: Darren Homrighausen, PhD

Overview

Key Ingredients

Let Y_t be the time series observed at time t

(In fMRI studies, the blood oxygenation level gets measured every 2s. Y_8 would be the 8^{th} s. measurement)

Mean Function:

$$\mu_t = \mathbb{E} Y_t$$

Autocovariance:

$$\gamma(s,t) = cov(Y_s, Y_t) = \mathbb{E}[(Y_s - \mu_s)(Y_t - \mu_t)]$$

(Hence, the variance at time t is $\gamma(t,t)$)

Example: Classical regression

Suppose $Y_t = \mu_t + \epsilon_t$

(Multiple regression would have $\mu_t = \beta_0 + X^\top \beta = \beta_0 + \sum_{j=1}^{p} x_j \beta_j$)

Suppose ϵ_t are all independent $N(0, \sigma^2)$

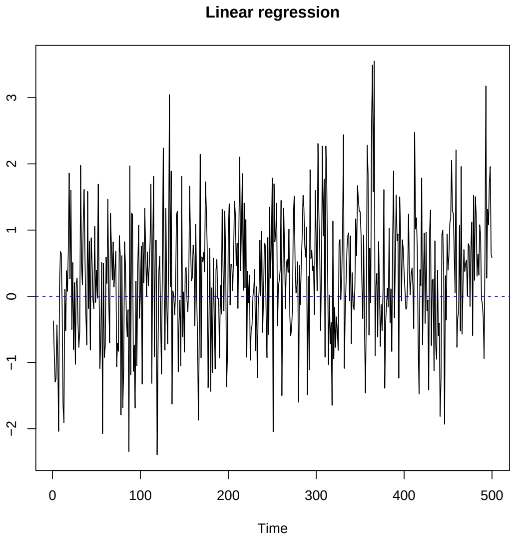
Then:

Mean Function:

$$\mathbb{E} Y_t = \mathbb{E}\mu_t + \mathbb{E}\epsilon_t = \mu_t$$

Autocovariance:

$$\gamma(s,t) = \mathbb{E}[(Y_s - \mu_s)(Y_t - \mu_t)] = \mathbb{E}\epsilon_s\epsilon_t = \begin{cases} \sigma^2 & \text{if } t = s \\ 0 & \text{if } t \neq s \end{cases}$$


Linear regression
3
2
1
0
−1
−2
0
100
200
300
400
500
Time

Example: Autoregression

Suppose $Y_t = 0.6Y_{t-1} + \epsilon_t$

Suppose ϵ_t are all independent $N(0, \sigma^2)$, and $Y_0 = 0$

Then:

Mean Function:

$$\mathbb{E}Y_t = \mathbb{E}[0.6Y_{t-1}] + \mathbb{E}\epsilon_t = 0$$

Autocovariance:

$$\gamma(s, t) = \mathbb{E}[(Y_s - \mu_s)(Y_t - \mu_t)] = \sigma^2 0.6^{|t-s|}$$

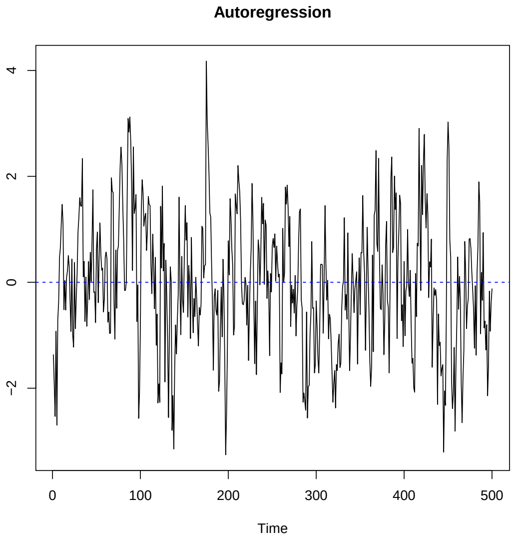
Autocovariance Function (ACF)

The autocovariance has the variance as a multiplier

$$\gamma(s,t) = \mathbb{E}[(Y_s - \mu_s)(Y_t - \mu_t)] = \sigma^2 0.6^{|t-s|}$$

So, we can just divide by this to get a scaled version

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}} = 0.6^{|t-s|}$$



Autoregression
4
2
0
−2
0
100
200
300
400
500
Time

White noise

The noise terms ϵ_t are known as white noise in time series

(The etymology is that the electromagnetic spectrum of white light looks like the analogous representation for ϵ_t)

This is a very important component to time series, as it represents disturbances that are unforecasteable

(Often, the whole sequence $(\epsilon_t)_{t=1}^{n}$ is known as a white noise process)

Regularity

A major topic in time series is stationarity

A time series $(Y_t)_{t=1}^{n}$ is stationary provided

- μ_t is constant over t
- The ACF depends on s, t only through $|t - s|$

Example:

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}} = 0.6^{|t-s|}$$

We might as well just write $s = t + h$ and write

$$\rho(h) = \frac{\gamma(h)}{\sqrt{\gamma(0)\gamma(0)}} = 0.6^{h}$$

Estimation and Testing

Under stationarity, we can estimate μ and $\rho(h)$ for $h = 0, 1, 2, \ldots$
(Using sample averages)

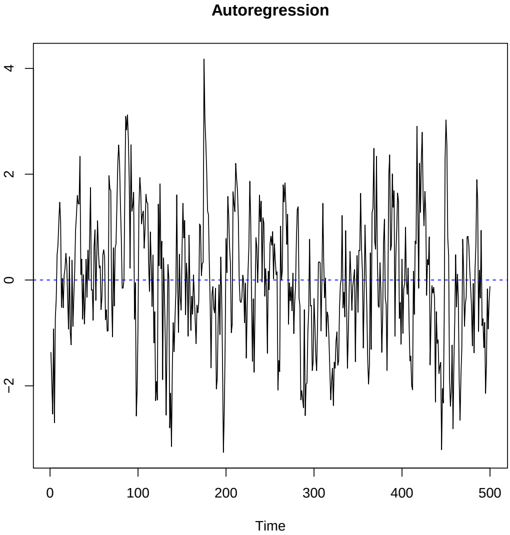
Write $\hat{\rho}(h)$ as the estimator of $\rho(h)$

Then $\hat{\rho}(h)$ has a particular distribution for white noise:

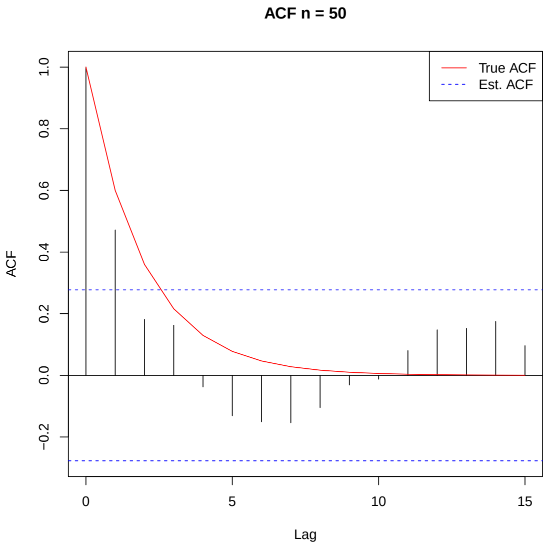
$$\hat{\rho}(h) \text{ is ``approx. distributed'' Normal}(0, 1/\sqrt{n})$$

Hence, if we get a lot of values outside of, say, $\pm 2/\sqrt{n}$, a white noise process is unlikely

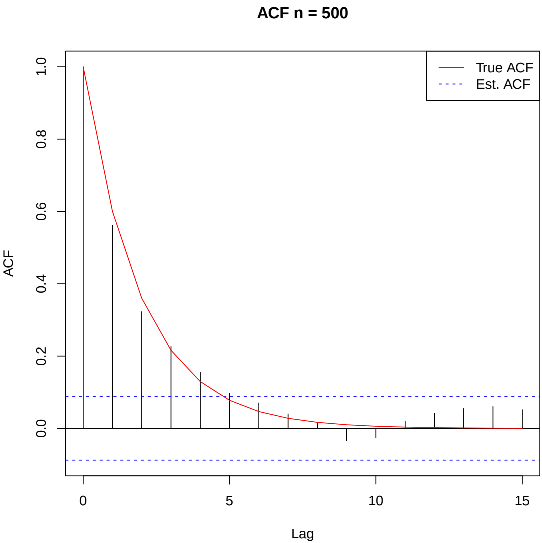
Data: autoregression



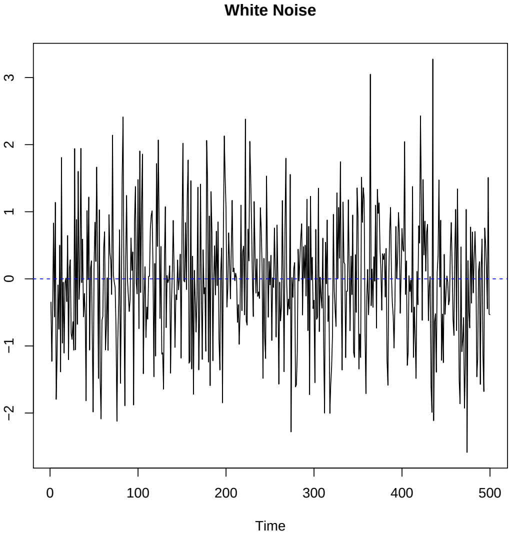
ACF: AUTOREGRESSION



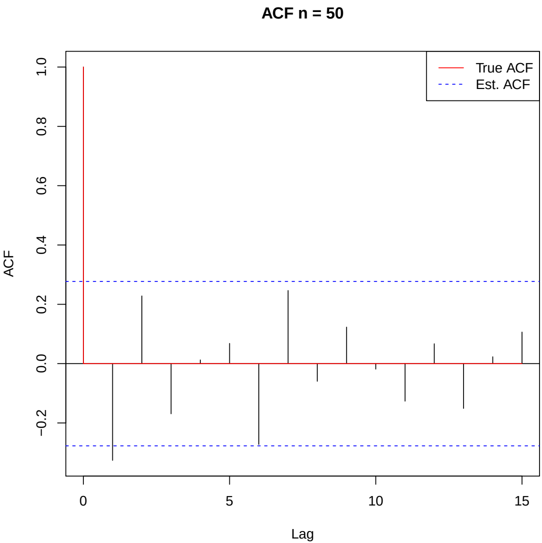
ACF: AUTOREGRESSION



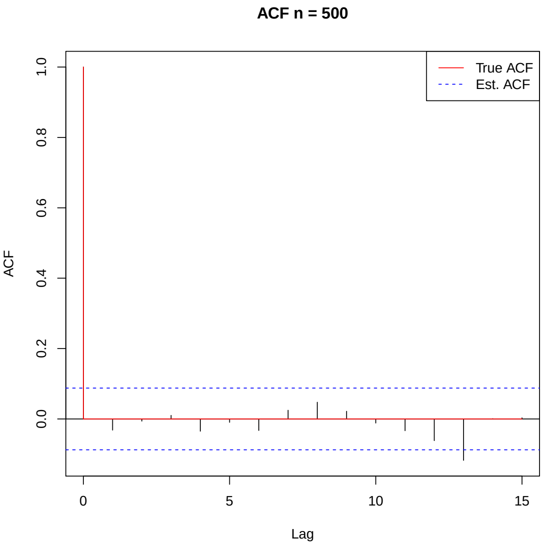
Data: white noise process



ACF: WHITE NOISE PROCESS



ACF: WHITE NOISE PROCESS



MCMC: An example

How does MCMC work?

The output of these iterations will look like

$$(Y_{mis}^1, (\mu^1, \Sigma^1)), (Y_{mis}^2, (\mu^2, \Sigma^2)), \ldots$$

This is a Markov chain

It has stationary distribution $p(Y_{mis}, (\mu, \Sigma) | Y_{obs})$

Assuming we

- Perform enough iterations
- Subselect from them iterations to break correlations
  (PROC MI has an autocorrelation plot)

Then we have independent draws of the missing observations from the posterior
(This is the goal of MCMC)

Why do MCMC?

We specify (and hence know)

- The likelihood $p(y|\theta)$
- The prior $p(\theta)$

Hence, via Bayes' theorem

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta)$$

The whole Bayesian bottleneck: We don't know $p(y)$!

(The aforementioned conjugate case being essentially the only exception)

Example

(To the SAS documentation: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mi_sect045.htm)

(Then to the R code: `MCMCstationaryDist.R`)