Estimating the ACF

STAT 1321/2320

Kiran Nihlani

Department of Statistics
University of Pittsburgh

Outline

Linear Process

ψ_j

An important class of stationary time series:

$$\boxed{X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}}$$

Time series is meaningful.

where $\{W_t\} \sim wn(0, \sigma_w^2)$ and μ, ψ_j are constants such that $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$.

$\mu_{xt} = E(X_t)$

- Mean function: $\mu_{xt} = \mu$ ✓
- Auto-covariance: $\gamma_x(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j-h}$ ✓

Linear Process - Examples

- White Noise: $\mu = 0$ and $\psi_j = 1$ for $j = 0$, and 0 otherwise. [hint: $X_t = W_t$]

$$X_t = \underline{\underline{W_t}} = \underline{\underline{\mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}}}$$

$$\mu = 0$$

$$\psi_0 = 1, \quad \psi_j = 0 \text{ for all } j \geq 1$$

$$X_t = 0 + 1 \cdot W_{t-j} + 0 \cdot$$

$$W_t$$

Linear Process - Examples

- White Noise: $\mu = 0$ and $\psi_j = 1$ for $j = 0$, and 0 otherwise. [hint: $X_t = W_t$]
- Moving Average, MA(1): $\psi_0 = 1, \psi_1 = \theta$, $\mu = 0$, $\psi_j = 0$ for $j \neq 1$ or 0.

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

AR(1) $\quad x_t = \phi x_{t-1} + w_t$.

$$X_t = \phi X_{t-1} + w_t$$

$$= \phi\left[\phi X_{t-2} + w_{t-1}\right] + w_t$$

$$= \phi^2 X_{t-2} + \phi w_{t-1} + w_t$$

$$= \phi^2\left[\phi X_{t-3} + w_{t-2}\right] + \phi w_{t-1} + w_t$$

$$= \phi^3 X_{t-3} + \underbrace{\phi^2 w_{t-2} + \phi w_{t-1} + w_t}$$

$$= \vdots$$

$$= w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \phi^3 w_{t-3} + \cdots$$

$$X_t = \sum_{j=0} \phi^j w_{t-j}$$

$\mu = 0$, $\psi_0 = 1$, $\psi_1 = \phi$, $\psi_2 = \phi^2$, ……

Linear Process - Examples

- White Noise: $\mu = 0$ and $\psi_j = 1$ for $j = 0$, and 0 otherwise. [hint: $X_t = W_t$]
- Moving Average, MA(1): $\psi_0 = 1, \psi_1 = \theta$

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

- Auto Regressive, AR(1): $\psi_0 = 1, \psi_1 = \phi, \psi_2 = \phi^2, \ldots$

$$\begin{aligned} X_t = \phi X_{t-1} + W_t = \phi[\phi X_{t-2} + W_{t-1}] + W_t &= \phi^2 X_{t-2} + [W_t + \phi W_{t-1}] \\ &= \phi^3 X_{t-3} + [W_t + \phi W_{t-1} + \phi^2 W_{t-2}] \\ &= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \phi^3 W_{t-3} + \cdots \end{aligned}$$

We need $|\phi| < 1$ for this to hold .(Why?)

Recall

Suppose that $\{X_t\}$ is a stationary time series.

- Its mean function is
$$\mu = E[X_t].$$
- Its autocovariance function is
$$\gamma(h) = Cov(X_{t+h}, X_t) = E[(X_{t+h} - \mu)(X_t - \mu)]$$
- Its autocorrelation function is
$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

Estimating the Functions

$E(\bar{X}) = \mu_{xt}$.

For observations $x_1, \ldots, x_n$ of a time series, the

- **sample mean** is $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$ (unbiased estimator - why?)

estimator

X_t , $E(X_t) = \mu_{xt}$.

sample :

$$E(\bar{x}) = E\left[\frac{1}{n}\sum_{t=1}^{n} x_t\right] = \frac{1}{n}\sum_{i=1}^{n} E(X_t)$$

$$= \frac{n\mu_{xt}}{n} = \mu_{xt}.$$

unbiasedness.

$-$ $\wedge$: sample.

Estimating the Functions

For observations $x_1, \ldots, x_n$ of a time series, the

- **sample mean** is $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$ (unbiased estimator - why?)
- **sample autocovariance function** is

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x}), \qquad \text{for } -n < h < n$$

Estimating the Functions

For observations $x_1, \ldots, x_n$ of a time series, the

- **sample mean** is $\bar{x} = \frac{1}{n} \sum_{t=1}^{n} x_t$ (unbiased estimator - why?)
- **sample autocovariance function** is

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \qquad \text{for } -n < h < n$$

- **sample autocorrelation function** is

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

Sample ACvF

$x_1, x_2, \ldots, x_n.$ $h \rightarrow$ $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-|h|} \end{pmatrix}$ h $\begin{pmatrix} x_{1+h} \\ x_{2+h} \\ x_{3+h} \\ \vdots \\ x_n. \end{pmatrix}$

Sample autocovariance function:

Biased.

$$\rightarrow \quad \hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} \underbrace{(x_{t+|h|} - \bar{x})(x_t - \bar{x})}_{n-|h|}, \qquad \text{for } -n < h < n.$$

This is the sample covariance of $(x_1, x_{h+1}), \ldots, (x_{n-h}, x_n)$, except that

- we subtract the full sample mean, but all n observation.
- we normalize by n instead of $n - h$.

  $a^T \hat{\gamma}(h)\, a \geq 0$.
  - This is to ensure the auto-covariance matrix is non-negative definite.
  - The estimator would be biased for either choice (n or $n - h$).

Sample ACF

fix sample size. → n

→ repeated samples.

→ statistic → sample AC.

↳ Dist of these statistics is the sampling dist.

Sample autocorrelation function:

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

The sample autocorrelation function has a sampling distribution that allows us to assess whether the data comes from a completely random or white series or whether correlations are statistically significant at some lags.

Large Sample Distribution of ACF

$n \to$ large.

Under general conditions*, if X_t is white noise, then for n large, the sample ACF, $\hat{\rho}_x(h)$, for $h = 1, 2, \ldots, H$, where H is fixed but arbitrary, is approximately normally distributed with zero mean and standard deviation given by

$$\sigma_{\hat{\rho}_x(h)} = \frac{1}{\sqrt{n}}$$

$$\hat{\rho}_x(h) \sim AN\left(0, 1/\sqrt{n}\right)$$

$$z_{\alpha/2} = 1.96.$$

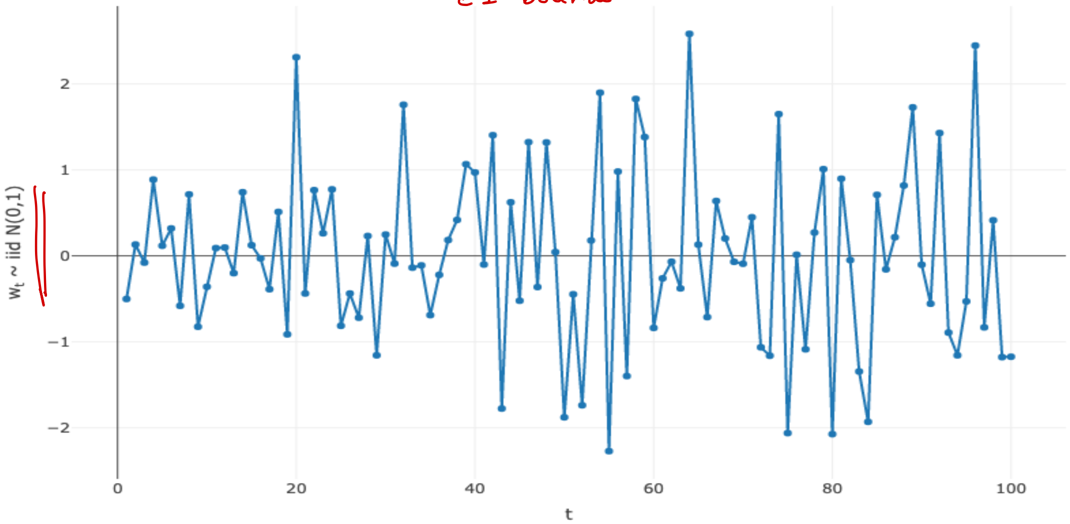
* The general conditions are that X_t is iid with finite fourth moment.

$$0 \pm z_{\alpha/2} \cdot 1/\sqrt{n}.$$

White Noise - Plot

simulating the series.
estimate the ACF and plot it.
CI bounds

$x_t = w_t$

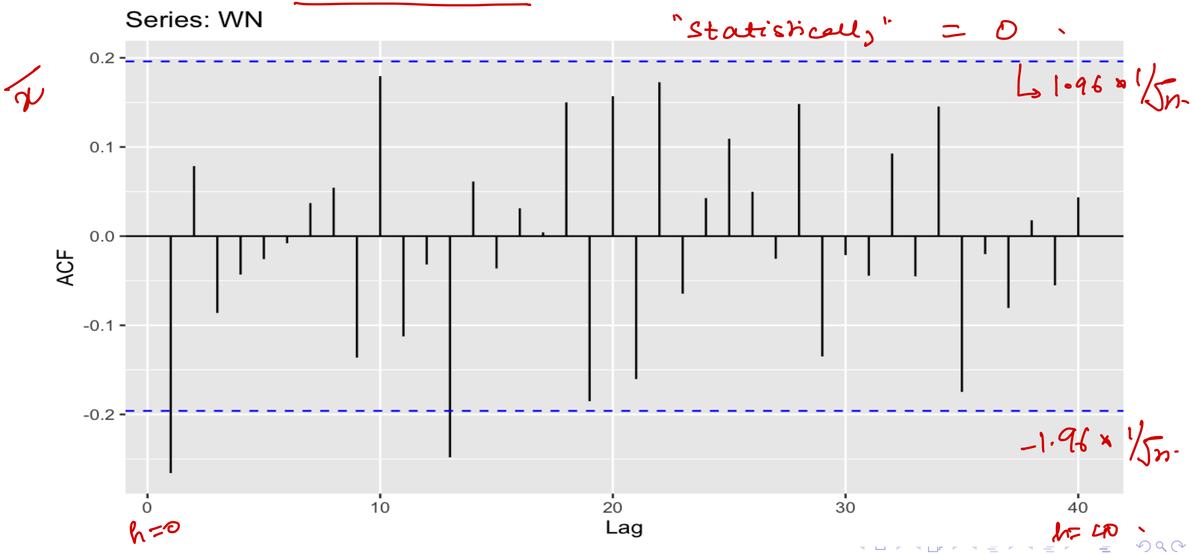


White Noise - Sample ACF

Theoretical ACFs. $c=95\%$.

$h=0$ $\rho_x(0) = 1$.

"statistically" $= 0$.



Observations - Sample ACF

- The plot is only plotted for positive lags. The ACF for $h = 0$ is omitted since it should be 1.

$0 \pm 1.96/\sqrt{n}$

- The blue dashed lines represent the 95% confidence interval bounds based on the asymptotic distribution.

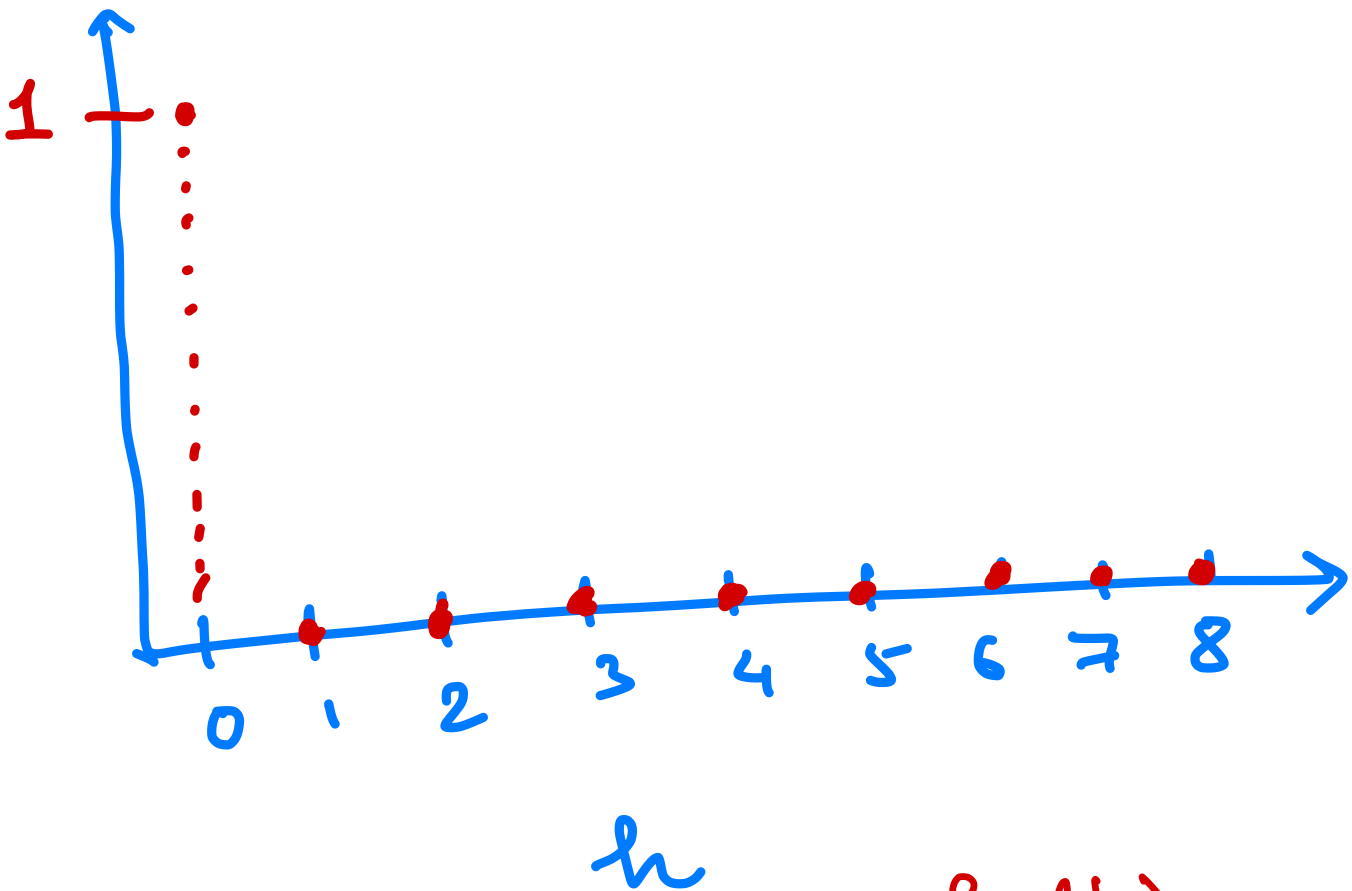
$\boxed{H_0: \rho(h) = 0}$

$\alpha = 0.05$

$c = 95\%.$

- Estimated ACF is insignificant for all lags except for $h = 1$ and $h = 13$. Implications?
    - Most correlations are zero (expected since it is IID white noise).
    - A 95% CI would imply that the correlation would be significant for 2 out of the 40 lags.
- Characteristic feature of the plot - ACF for 95% of the lags is 0 (insignificant).

Theoretical ACF plot $\quad x_t = w_t, \; w_t \sim wn(0, \sigma^2)$



$$\rho_x(h) = \begin{cases} 1 & h = 0 \\ 0 & h \neq 0 \end{cases}$$

Sample ACF in R

- Basic

  acf ( )

  plots ACF starting at lag, h = 0

- GG Plot

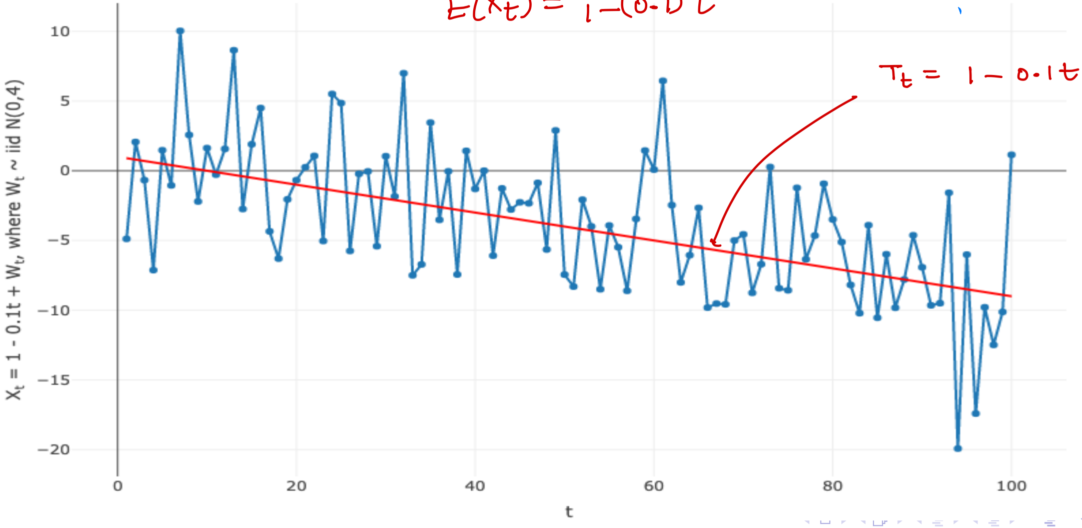
  ggAcf ( )

  plots ACF starting at lag, h = 1

Trend Model - Plot

$$X_t = \underbrace{1 - (0.1)t}_{T_t} + \underbrace{w_t}_{I_t}, \quad w_t \sim N(0, 4)$$

$$E(X_t) = 1 - (0.1)t$$



$$\hat{\gamma}_x(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x})$$

Sample cov. between

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-|h|} \end{pmatrix} \text{ and } \begin{pmatrix} x_{1+h} \\ x_{2+h} \\ \vdots \\ x_n \end{pmatrix}$$

Shifted by lag "h"

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

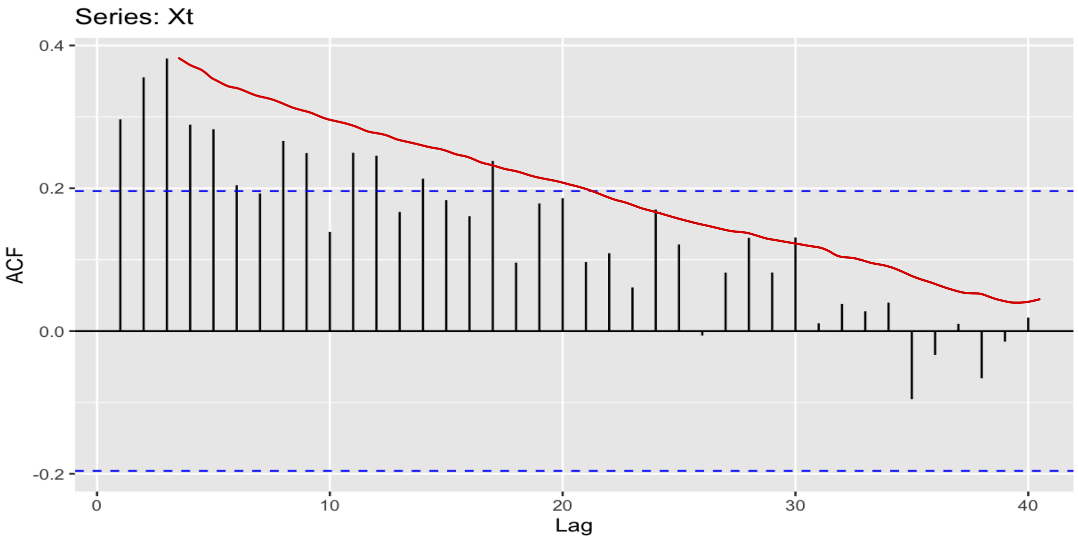
When lag is small :

When lag increases :

$$h = 1 \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

Autocovariance ↑

Trend Model - Sample ACF



Observations - Sample ACF for Trend Model

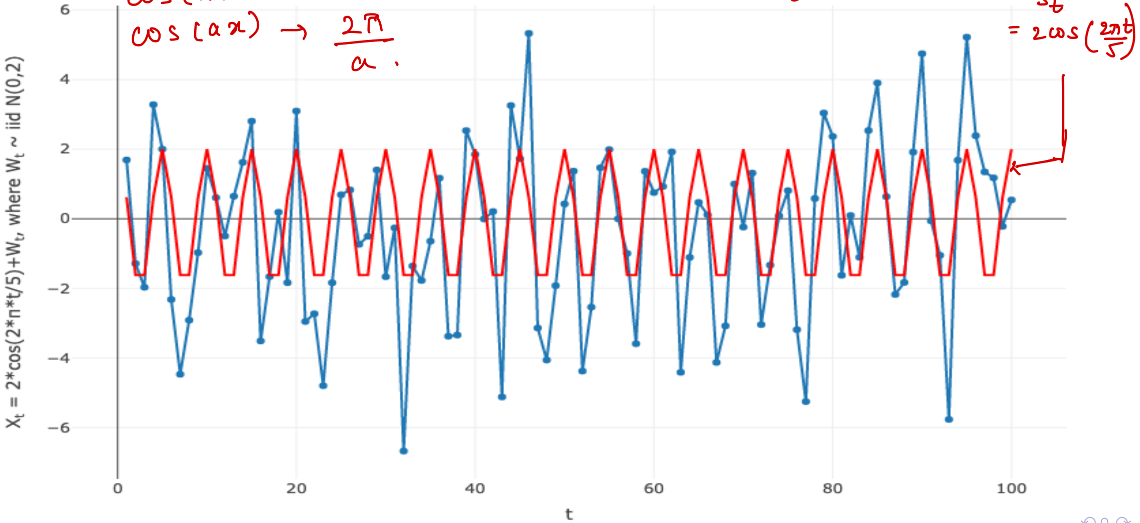
- Autocorrelations for small lags are large and positive because observations nearby in time are also nearby in value.
- ACF slowly decreases as the lags increase.
- Characteristic feature of the plot - slow decay.

Seasonal Model - Plot

$$X_t = \underbrace{2\cos\left(\frac{2\pi t}{5}\right)}_{S_t} + \underbrace{w_t}_{I_t}.$$

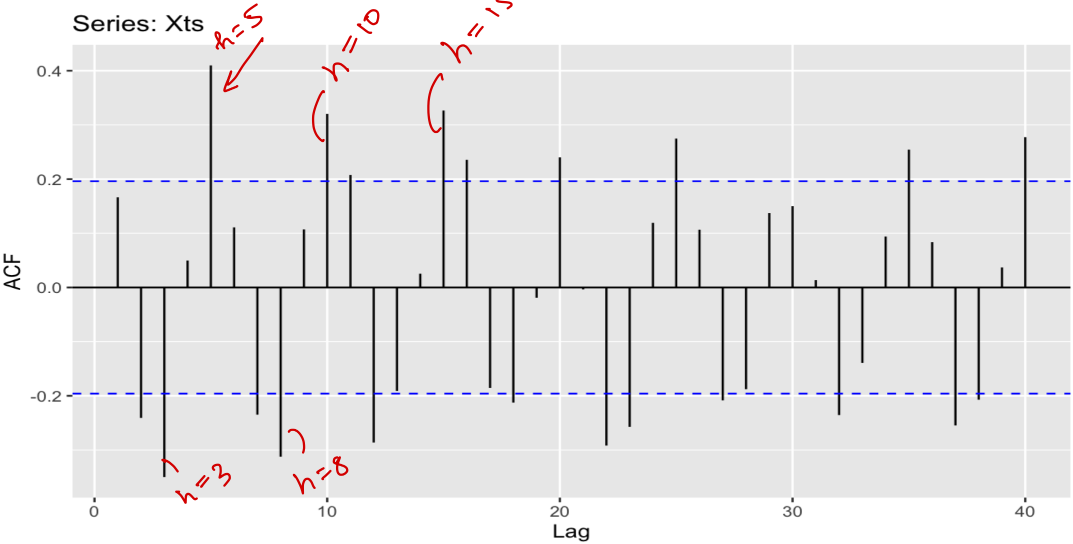
$\cos(x) \rightarrow 2\pi$

$\cos(ax) \rightarrow \frac{2\pi}{a}.$



Seasonal Model - Sample ACF

5, multiples of 5.



Observations - Sample ACF for Seasonal Model

- The series was simulated from

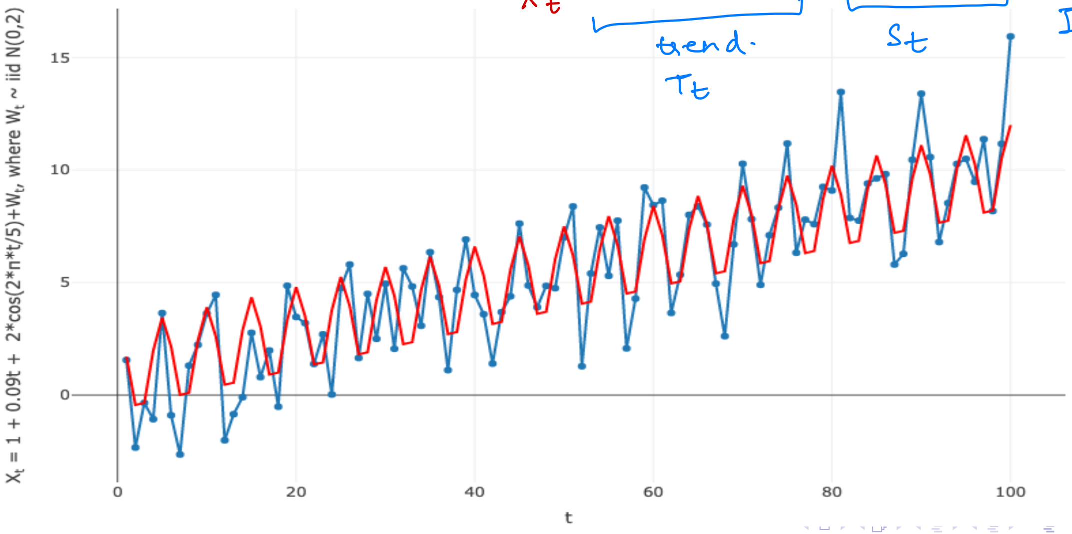
$$X_t = 2cos(2\pi t/5) + W_t \qquad \text{where } W_t \sim N(0, \sigma_w = 2)$$

- The period of the series is 5.
- The autocorrelations are larger for the seasonal lags (at multiples of the seasonal period) than for other lags.
    - Significant positive correlations at lags, $h = 5, 10, 15, \ldots$.
    - Significant negative correlations that are periodic as well.
- Characteristic feature of the plot - periodic.

seasonal lags or multiples of them.

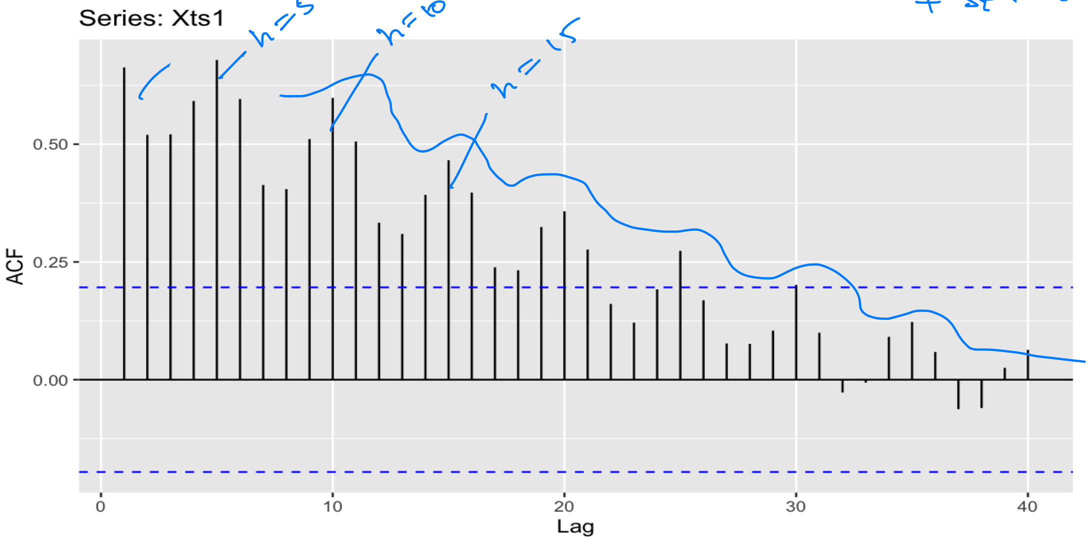
Seasonal + Trend Model - Plot

$$X_t = \underbrace{1 + 0.09t}_{\text{trend. } T_t} + \underbrace{2\cos\left(\frac{2\pi t}{5}\right)}_{S_t} + \underbrace{w_t}_{I_t}$$



Seasonal + Trend Model - Sample ACF

$X_t = 1 + 0.09t + S_t + I_t$.



Observations - Seasonal + Trend Model

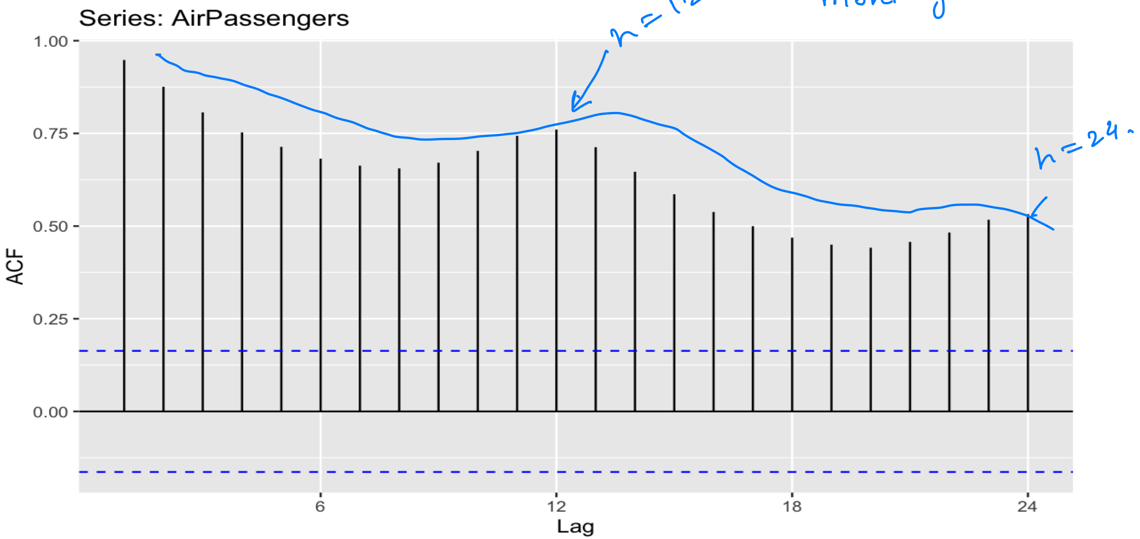
- The series was simulated from

$$X_t = 1 + 0.09t + 2cos(2\pi t/5) + W_t \qquad \text{where } W_t \sim N(0, \sigma_w = 2)$$

- Positive significant correlations for smaller lags that slowly decrease for larger lags.
- There is still some periodicity in the ACF values.
- Characteristic Feature - Slow decay along with periodicity.

AirPassengers Data - Sample ACF

12
monthly data



Notes

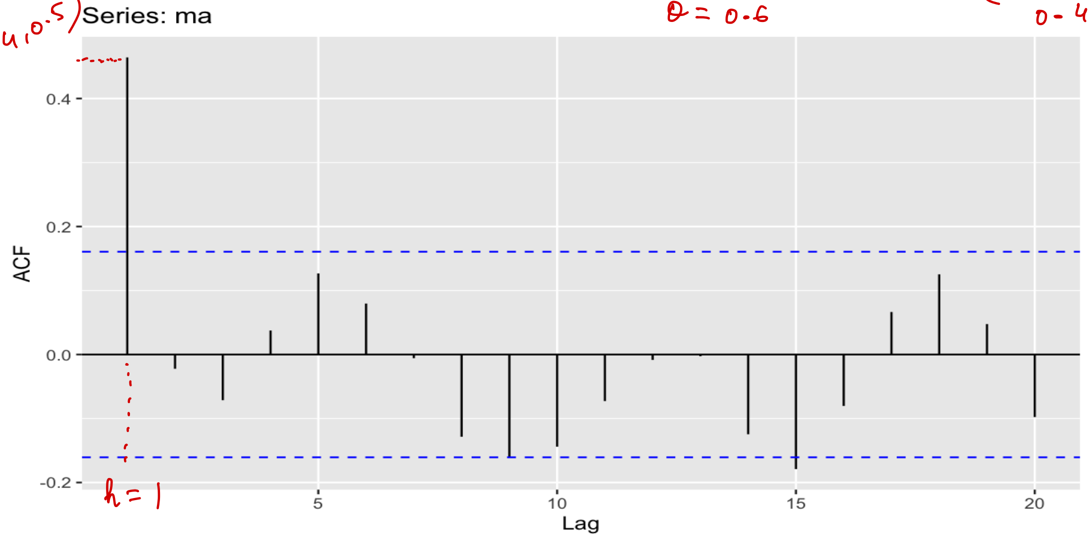
- Little can be inferred from an ACF plot where trend dominates all other features.
  - It may be useful to detrend the series before calculating the ACFs.
  - If trend is of primary interest, then it should be modelled, rather than removed and ACF plot may not be as helpful.
- If the seasonal variation is removed from seasonal data, then the ACF plot may provide useful information about correlation in the short term.

Moving Average (1) - Sample ACF

$\theta / 1+\theta^2 = \frac{0.6}{1+(0.6)^2} = 0.44$

$\theta = 0.6$

$(0.4, 0.5)$



$h = 1$

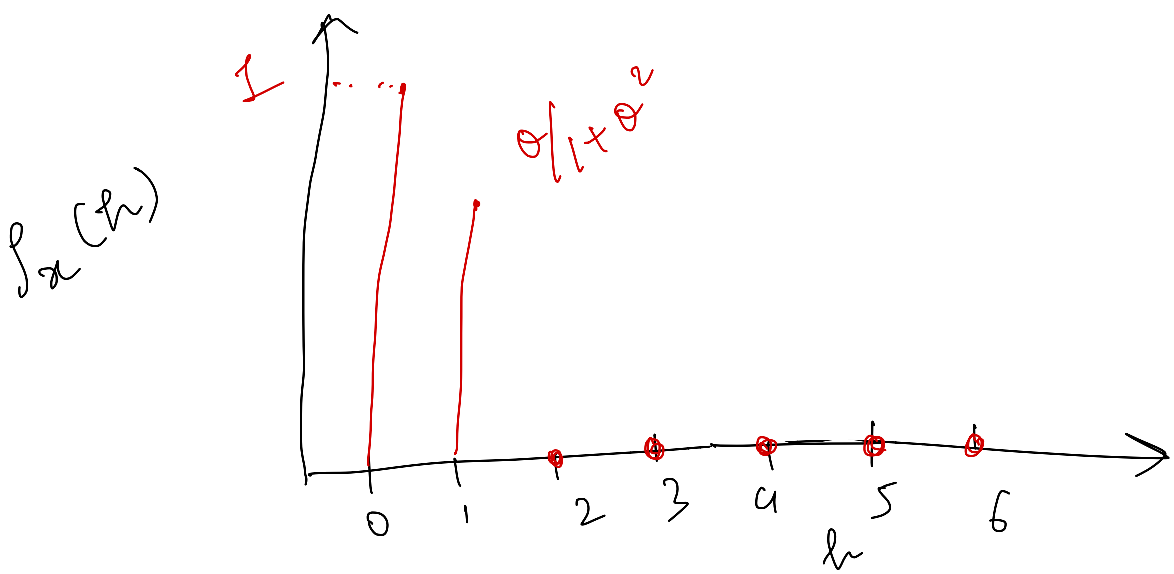
MA(1) : Theoretical ACF

$$X_t = w_t + \theta w_{t-1}$$

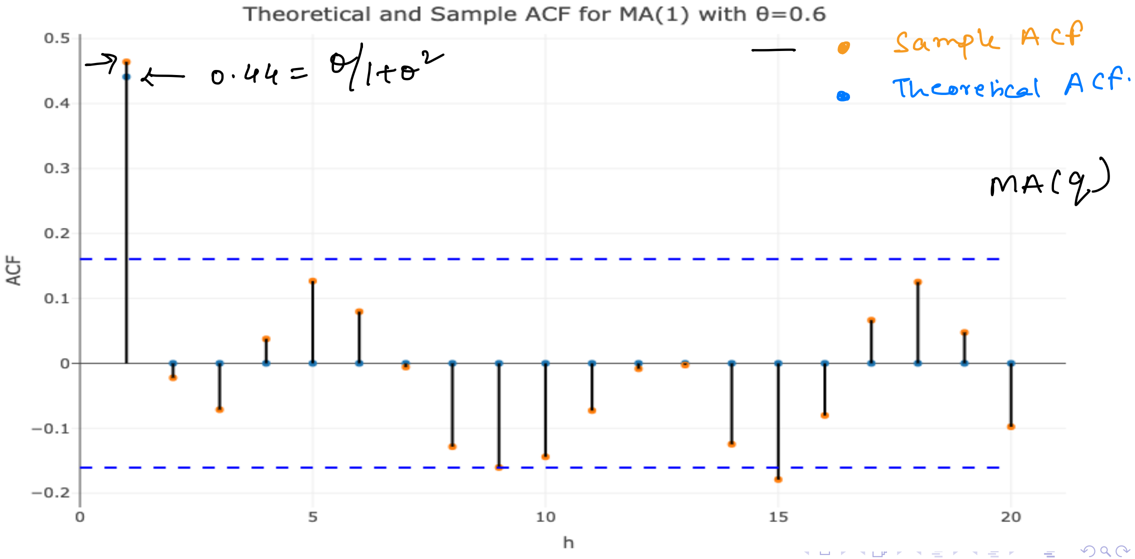
$$\gamma_x(h) = \begin{cases} (1+\theta^2)\sigma^2 & h = 0 \\ \theta\sigma^2 & h = 1 \\ 0 & h > 1 \end{cases}$$

$$\rho_x(h) = \begin{cases} 1 & h=0 \\ \frac{\theta}{1+\theta^2} & |h|=1 \\ 0 & |h|>1 \end{cases}$$

$\theta > 0$



Moving Average (1) - Theoretical vs Sample ACF



Observations

* Z_t should be w_t (white noise)

- The data is simulated from the following MA(1) model

$X_t = w_t + \theta w_{t-1}$ $\qquad$ $w_t \sim N(0,1)$

$$X_t = Z_t + \theta Z_{t-1} \qquad \text{where } \theta = 0.6, Z_t \sim N(0, 1)$$

- Theoretically, the ACF is equal to $\theta/(1 + \theta^2)$ for $h = 1$ and zero for $h \geq 2$.
- The sample ACF supports this as we have statistically insignificant correlation values at lags greater than 1.
- Legend for the plot on previous slide:
  - Blue dots give the theoretical ACF.
  - Orange dots (connected by black solid lines) give the sample ACF.
  - Blue dashed lines are the 95% (asymptotic) confidence interval bounds.

Moving Average, MA(q) - Model

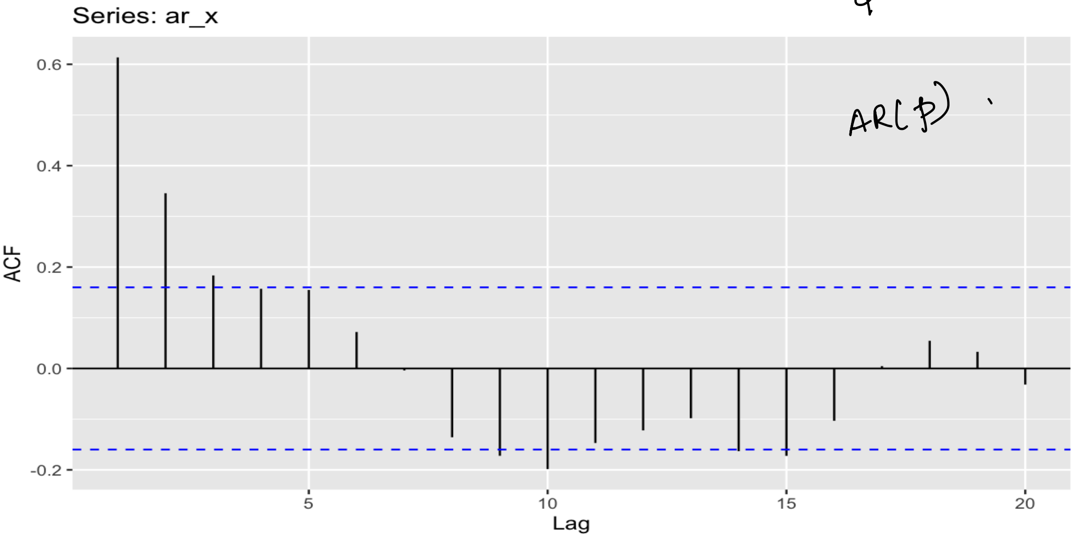
ACF

- An MA(q) model would theoretically have non-zero ACF until lag q and zero for $|h| > q$.
- The theoretical ACF values may be positive or negative depending on the value of θ.
- Characteristic feature of sample ACF - Significant correlations for $h \leq q$ and close to zero or non-significant thereafter.

$$X_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}.$$

Autoregressive, AR(1) Model - Sample ACF

$\phi = 0.6$.

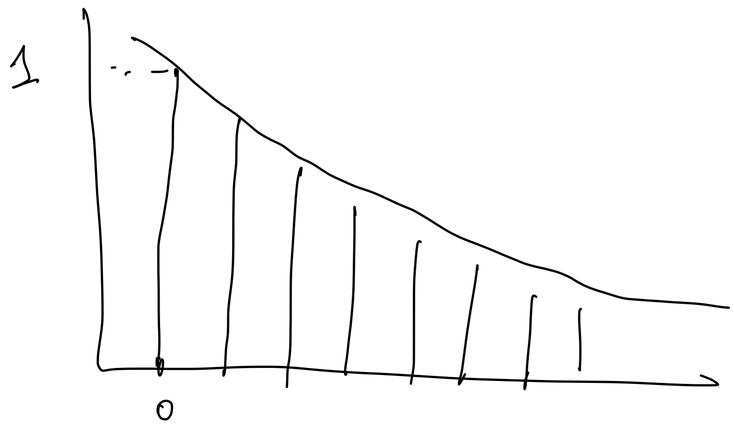


$AR(p)$.

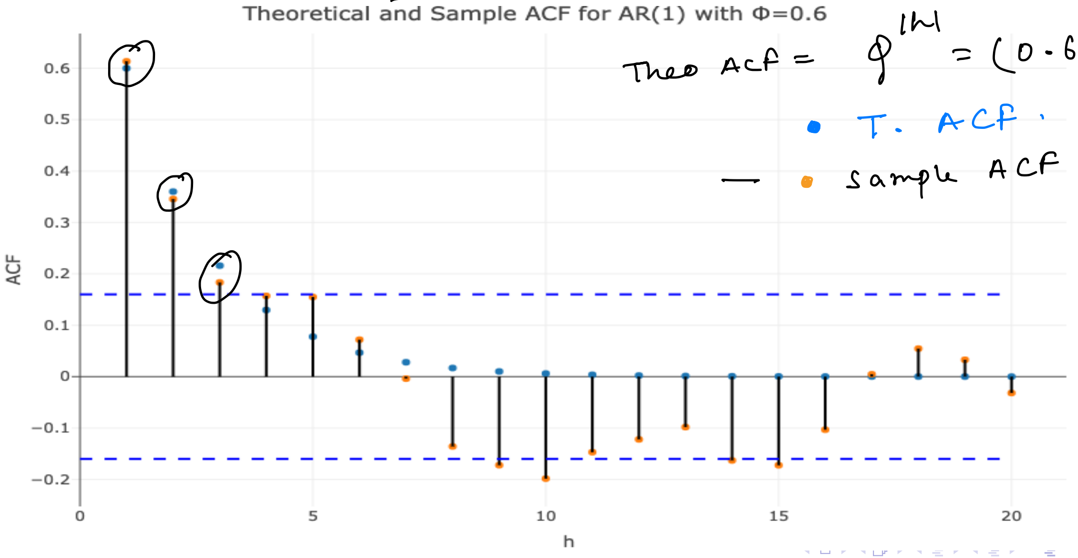
AR(1)

$$x_t = \phi\, x_{t-1} + w_t .$$

$$\rho_x(h) = \phi^{|h|} \qquad |\phi| < 1$$



Autoregressive (1) - Theoretical vs Sample ACF



Theo ACF $= \phi^{|h|} = (0.6)^h$

• T. ACF.

— • sample ACF.

partial auto-correlation.

AR(p)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + w_t$$

PACF

partial correlation.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 .$$

cond^n correlations.

$cor(y, x_2 | x_1)$

$cor(y, x_1 | x_2)$

partial correlation.

Observations

* Z_t should be W_t

- The data is simulated from the following MA(1) model

$$X_t = \phi X_{t-1} + \overset{W_t}{Z_t} \qquad \text{where } \phi = 0.6, \overset{W_t}{Z_t} \sim N(0, 1)$$

- Theoretically, the ACF is equal to $\phi^{|h|}$ for $h \geq 1$.
- Characteristic feature - exponential decay to 0.

ACF plot is not enough.

Need PACF.

ACF and Model Uniqueness

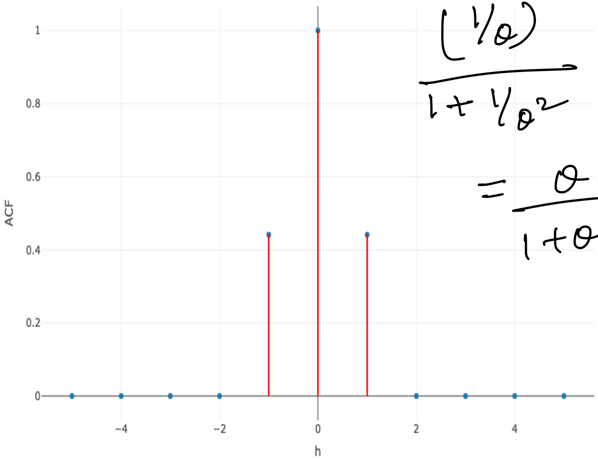
$$\rho_{x_A}(h) = \begin{cases} 1 & h=0 \\ \theta/1+\theta^2 & |h|=1 \\ 0 & o.w. \end{cases}$$

- Consider the following first order MA processes: MA(1).

$$A: \quad X_t = \overset{w_t}{Z_t} + \theta \overset{w_{t-1}}{Z_{t-1}}$$

$$B: \quad X_t = \underset{w_t}{Z_t} + \frac{1}{\theta} \underset{w_{t-1}}{Z_{t-1}}$$

- It can easily be shown that these two different processes have exactly the same ACF.
- See figure on right for $\theta = 0.6$.



$$\frac{(1/\theta)}{1 + 1/\theta^2} = \frac{\theta}{1+\theta^2}$$

* Z_t should be W_t

ACF and Model Identification

/ correlation

- A given stochastic process has a unique covariance structure but the converse is not true in general.

" Invertibility "

- The ACF does not uniquely identify the underlying model.
- Sample ACF can therefore be used to make educated guess for models but we can't be sure.
- Further, the sample ACF will rarely fit a perfect theoretical pattern. A lot of the time you just have to try a few models to see what fits.