Stat 565

The PACF & Estimation for ARMA

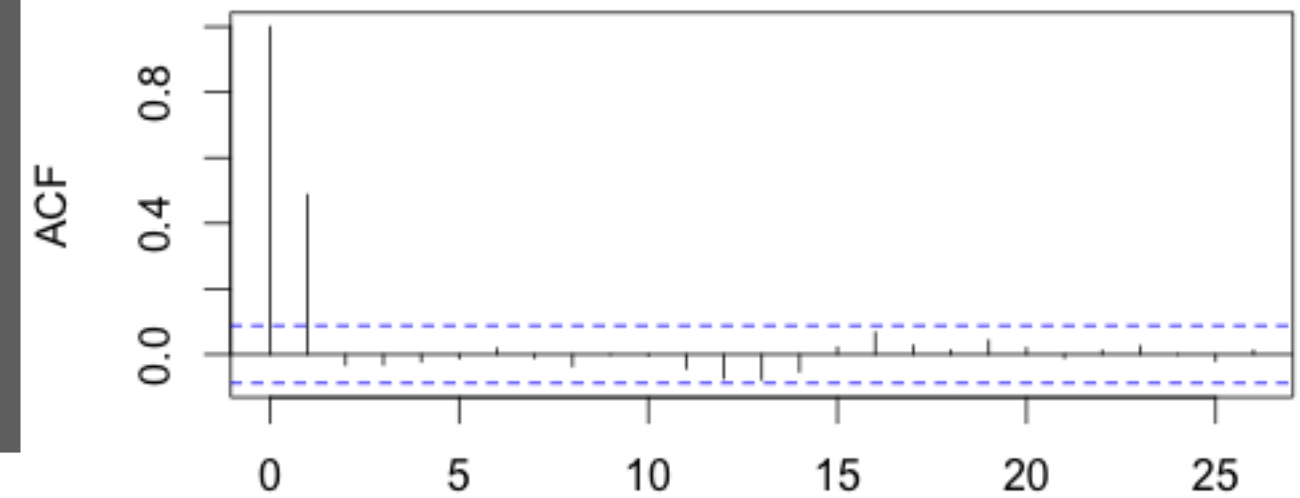
Jan 26 2016

Charlotte Wickham

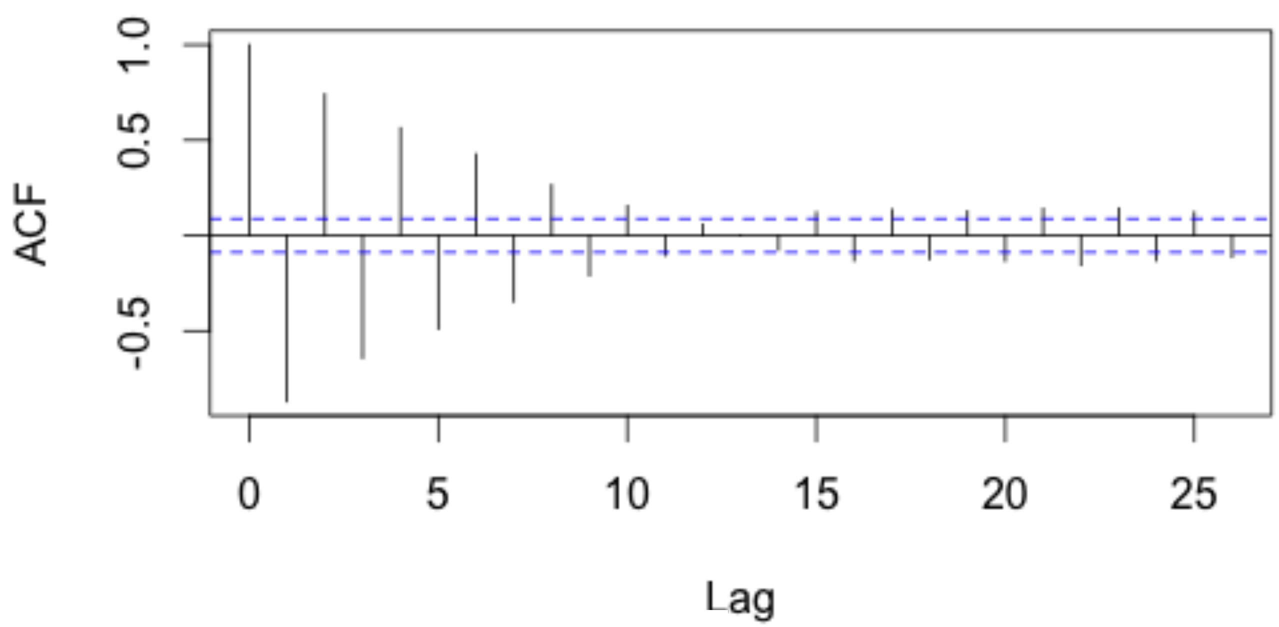
stat565.cwick.co.nz

Which **basic** models might these simulated data come from?
AR(1), MA(1), or white noise

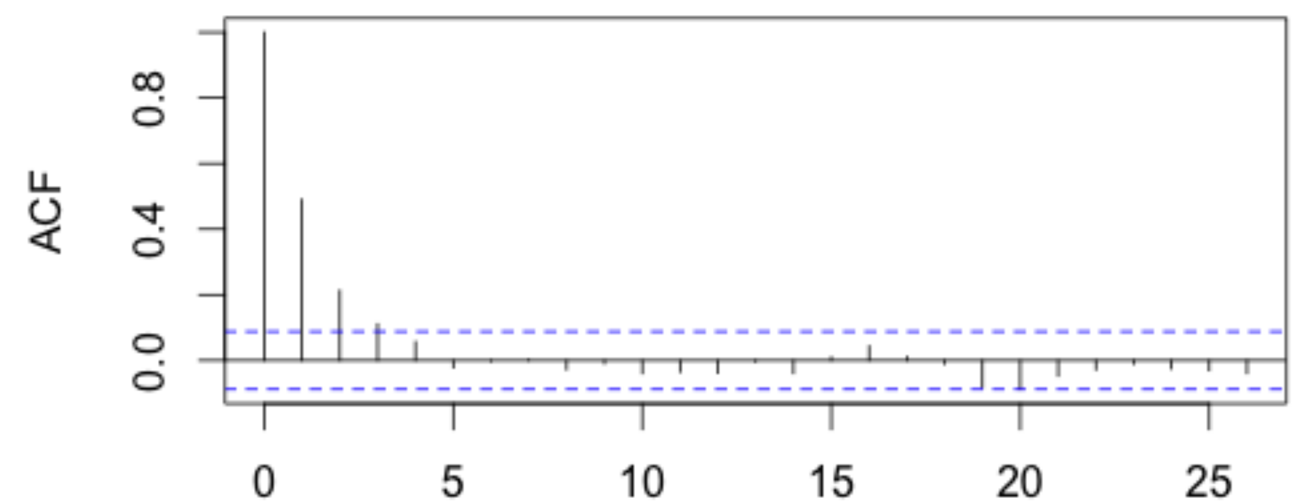
1



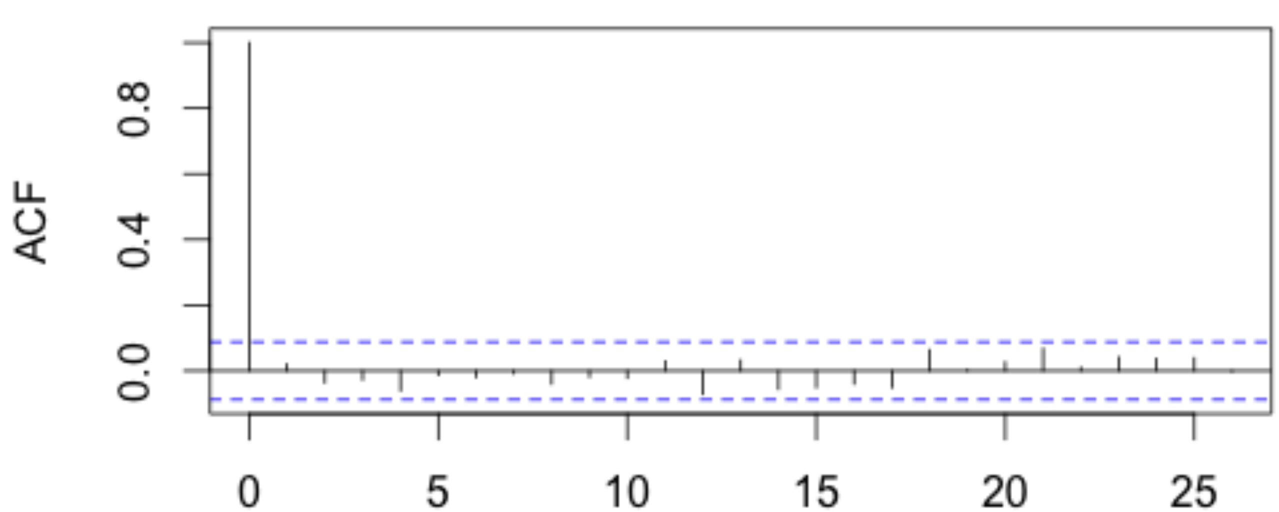
2



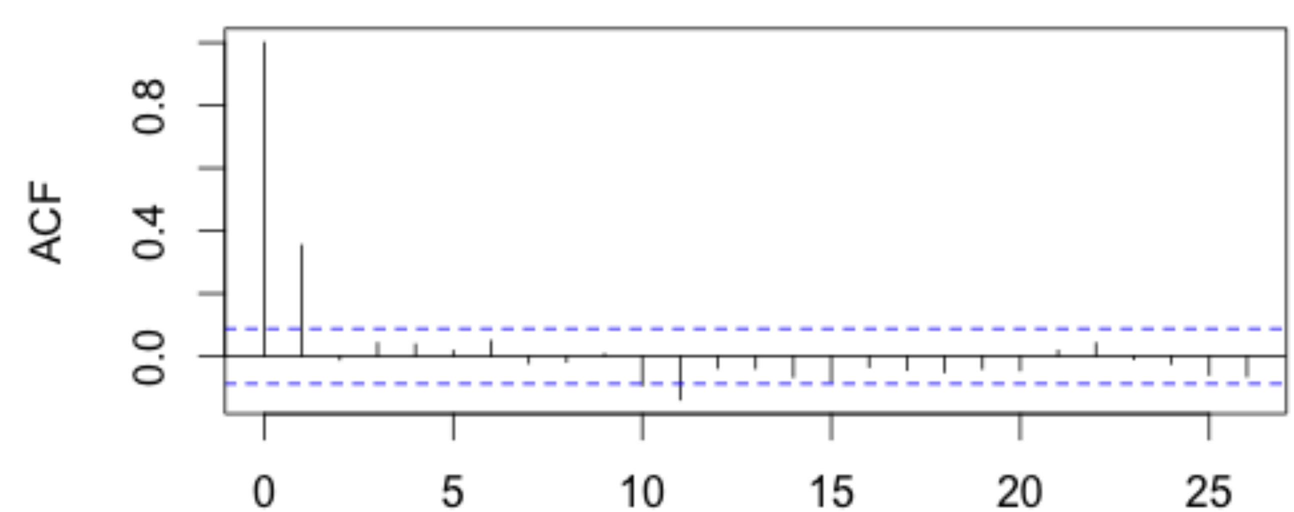
3



4



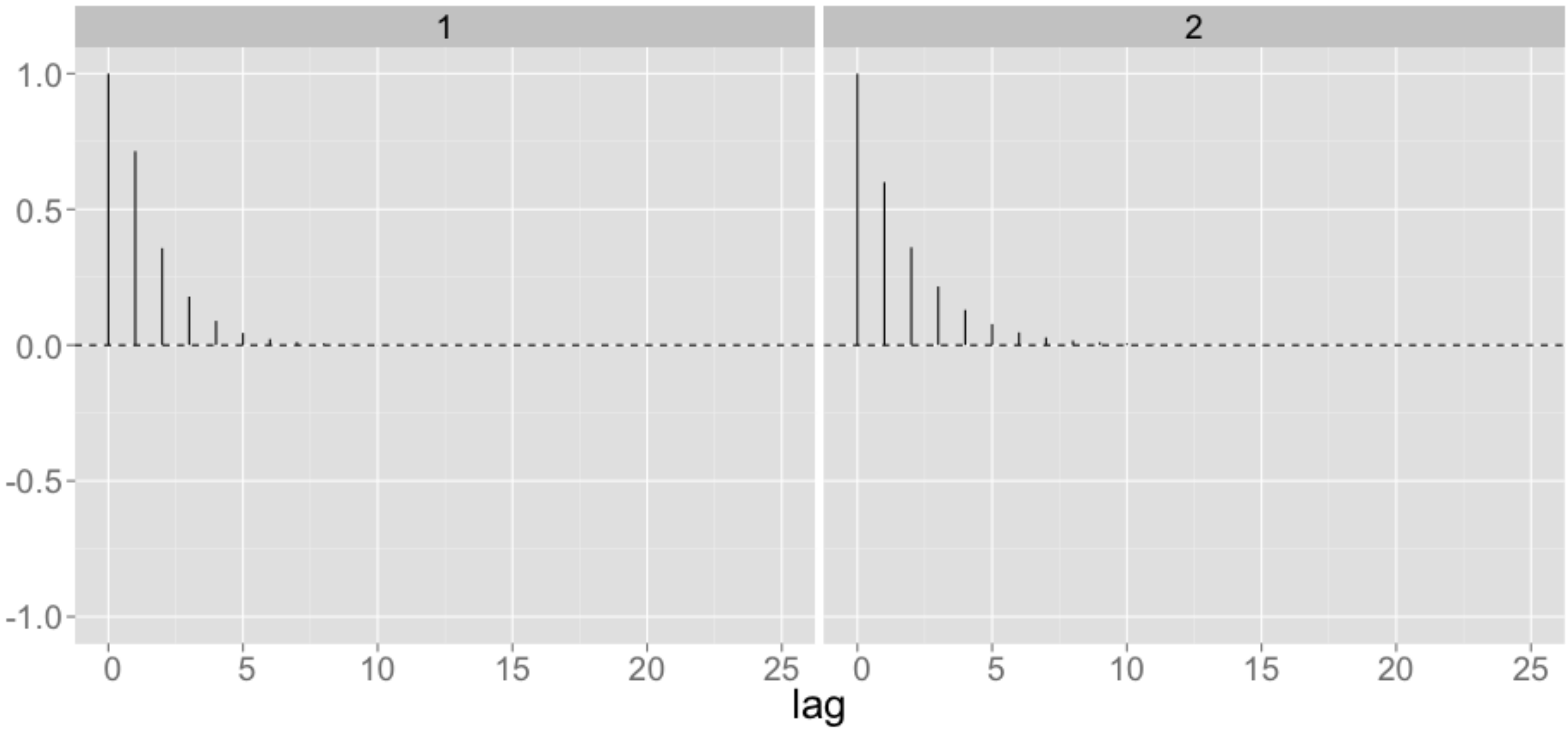
5



ACF

One is AR(1), $\alpha_1 = 0.6$

The other is ARMA(1, 1), $\beta_1 = 0.5$, $\alpha_1 = 0.5$



Which is which?

Partial autocorrelation function

Basic idea: what is the correlation between x_t and x_{t+h}, after taking into account x_{t+1}, x_{t+2}, ..., x_{t+h-1}?

Technically:

Regress x_t on x_{t+1}, x_{t+2}, ..., x_{t+h-1} to find the fitted value $\hat{x}_t$.

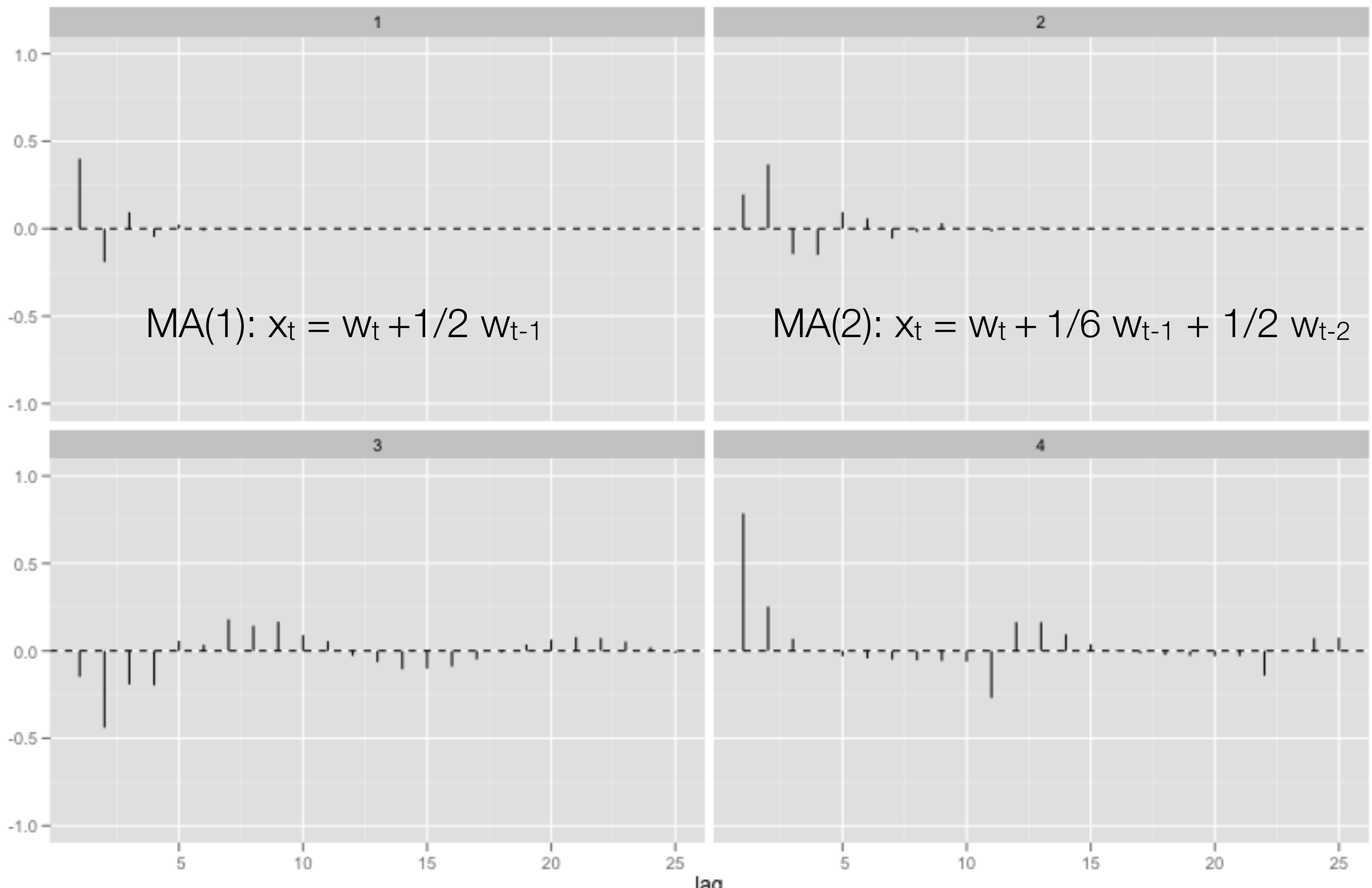
Regress x_{t+h} on x_{t+1}, x_{t+2}, ..., x_{t+h-1} to find the fitted value $\hat{x}_{t+h}$.

Find $\mathrm{cor}(x_t - \hat{x}_t, x_{t+h} - \hat{x}_{t+h})$, call this $\mathrm{PACF}(h) = \phi_{hh}$

Section 3.4 S&S for more detail

PACF

MA(q)



MA(1): $x_t = w_t + 1/2\, w_{t-1}$

MA(2): $x_t = w_t + 1/6\, w_{t-1} + 1/2\, w_{t-2}$

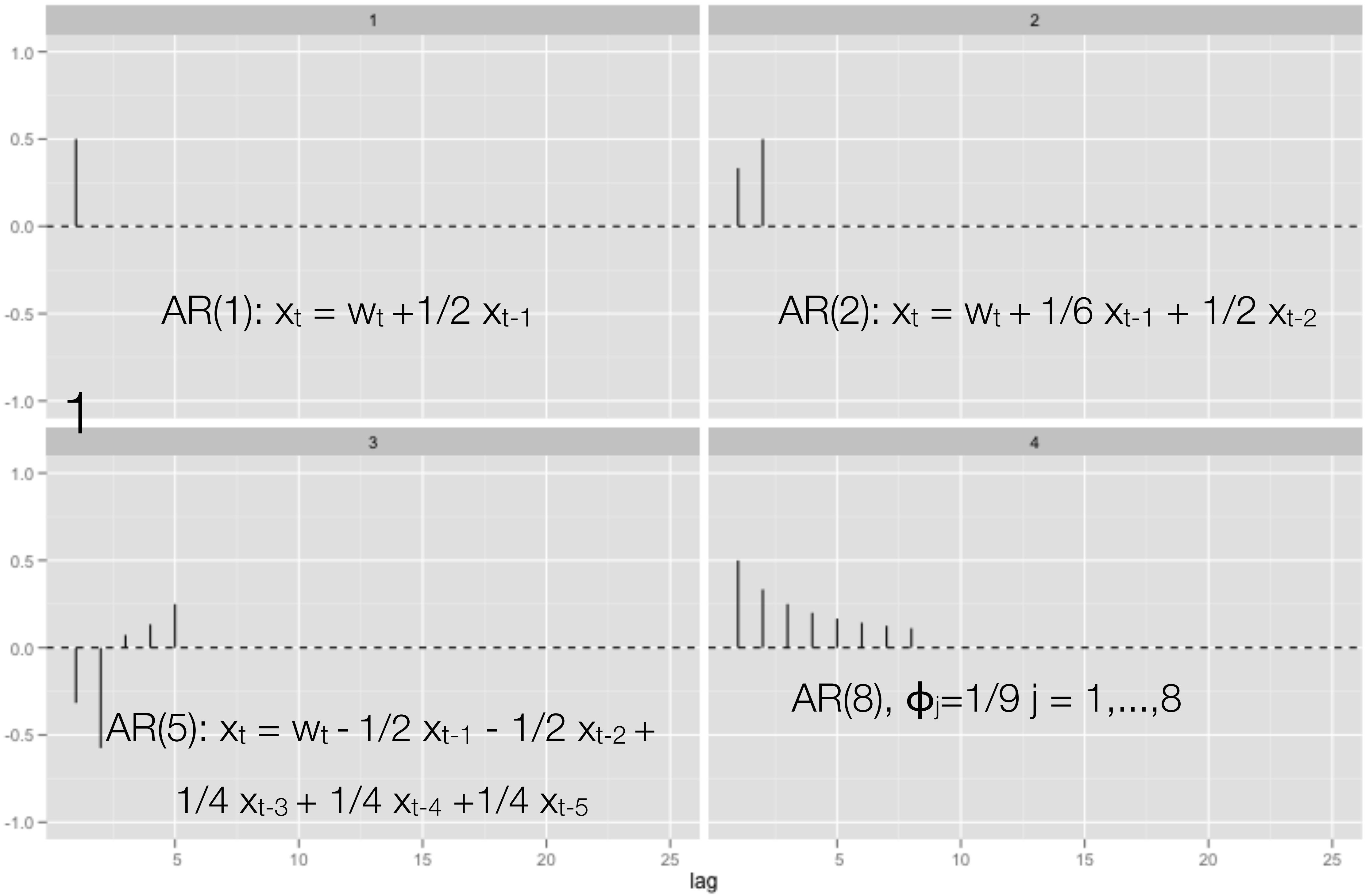
MA(5): $x_t = w_t - 1/2\, w_{t-1} - 1/2\, w_{t-2} +$

$1/4\, w_{t-3} + 1/4\, w_{t-4} + 1/4\, w_{t-5}$

MA(10), $\theta_j = 1/2$ $j = 1,\ldots,10$

PACF

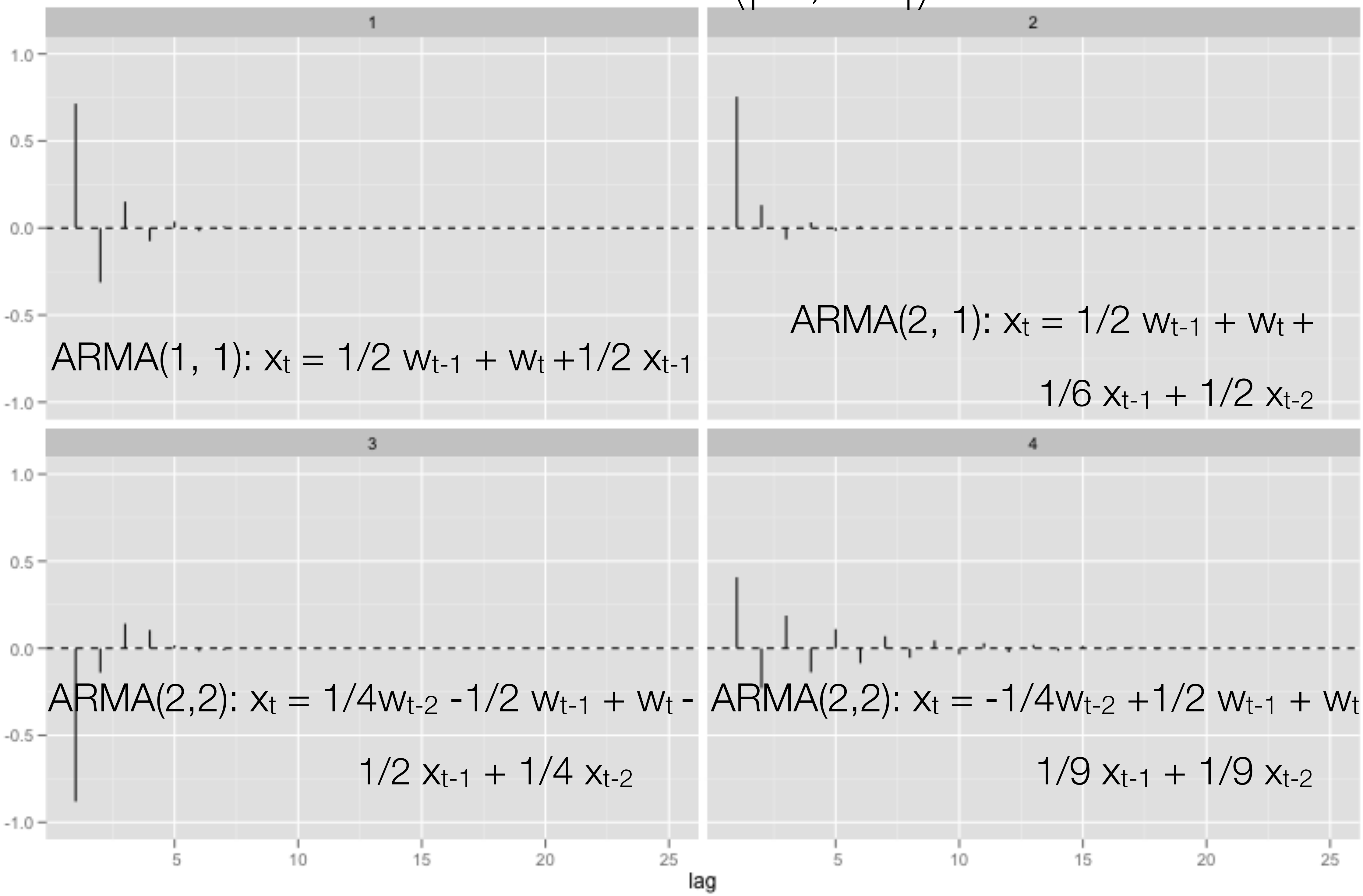
AR(p)



1
2
3
4
1.0
0.5
0.0
-0.5
-1.0
5
10
15
20
25
lag
AR(1): $x_t = w_t + 1/2\ x_{t-1}$
AR(2): $x_t = w_t + 1/6\ x_{t-1} + 1/2\ x_{t-2}$
AR(5): $x_t = w_t - 1/2\ x_{t-1} - 1/2\ x_{t-2} +$
$1/4\ x_{t-3} + 1/4\ x_{t-4} + 1/4\ x_{t-5}$
AR(8), $\phi_j = 1/9$ $j = 1,\ldots,8$

PACF

ARMA(p, q)



ARMA(1, 1): $x_t = 1/2\ w_{t-1} + w_t + 1/2\ x_{t-1}$

ARMA(2, 1): $x_t = 1/2\ w_{t-1} + w_t + 1/6\ x_{t-1} + 1/2\ x_{t-2}$

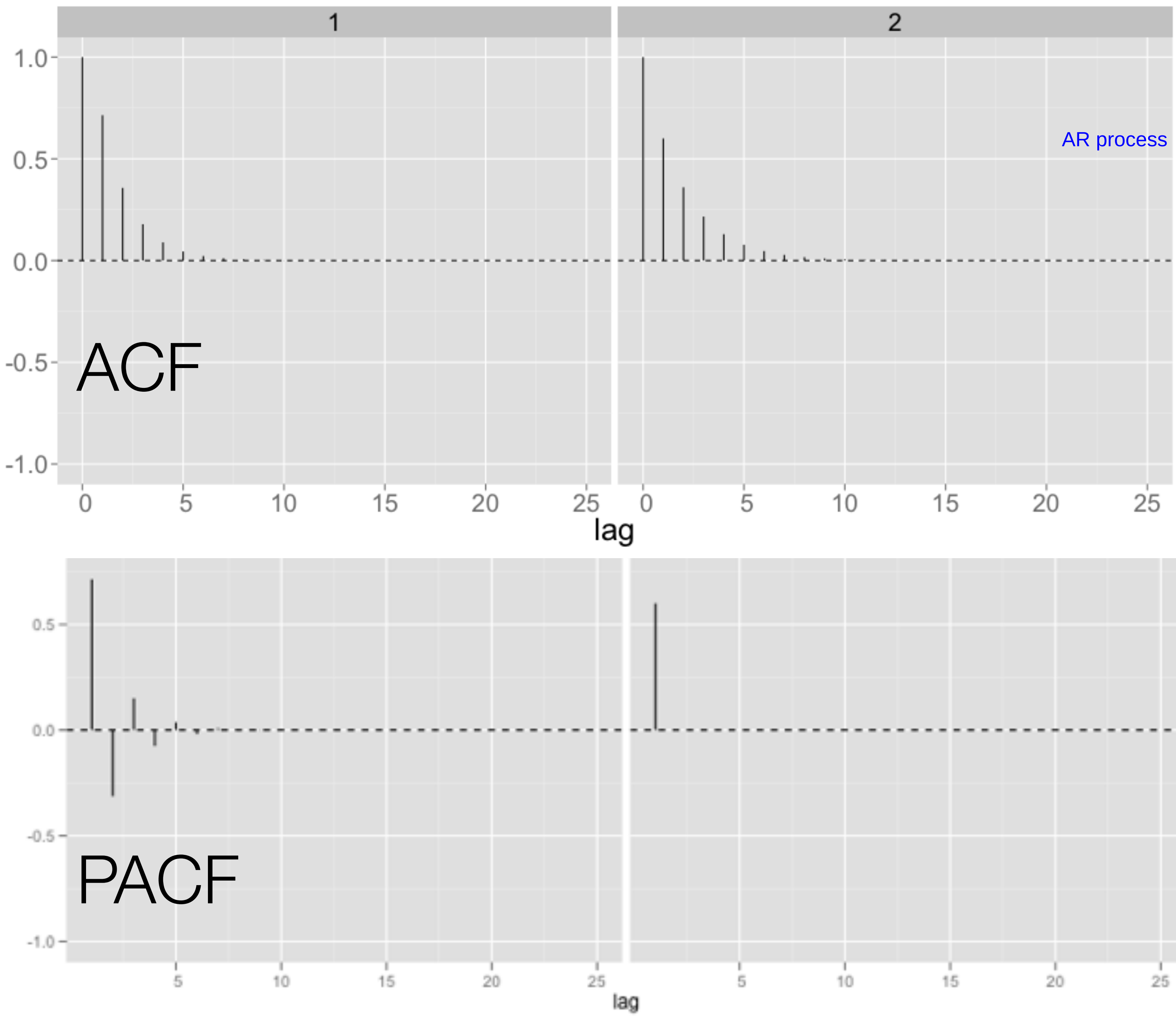
ARMA(2,2): $x_t = 1/4 w_{t-2} - 1/2\ w_{t-1} + w_t - 1/2\ x_{t-1} + 1/4\ x_{t-2}$

ARMA(2,2): $x_t = -1/4 w_{t-2} + 1/2\ w_{t-1} + w_t + 1/9\ x_{t-1} + 1/9\ x_{t-2}$

	MA(q)	AR(p)	ARMA(p, q)
ACF	zero lags > q	tails off	tails off
PACF	tails off	zero lags > p	tails off

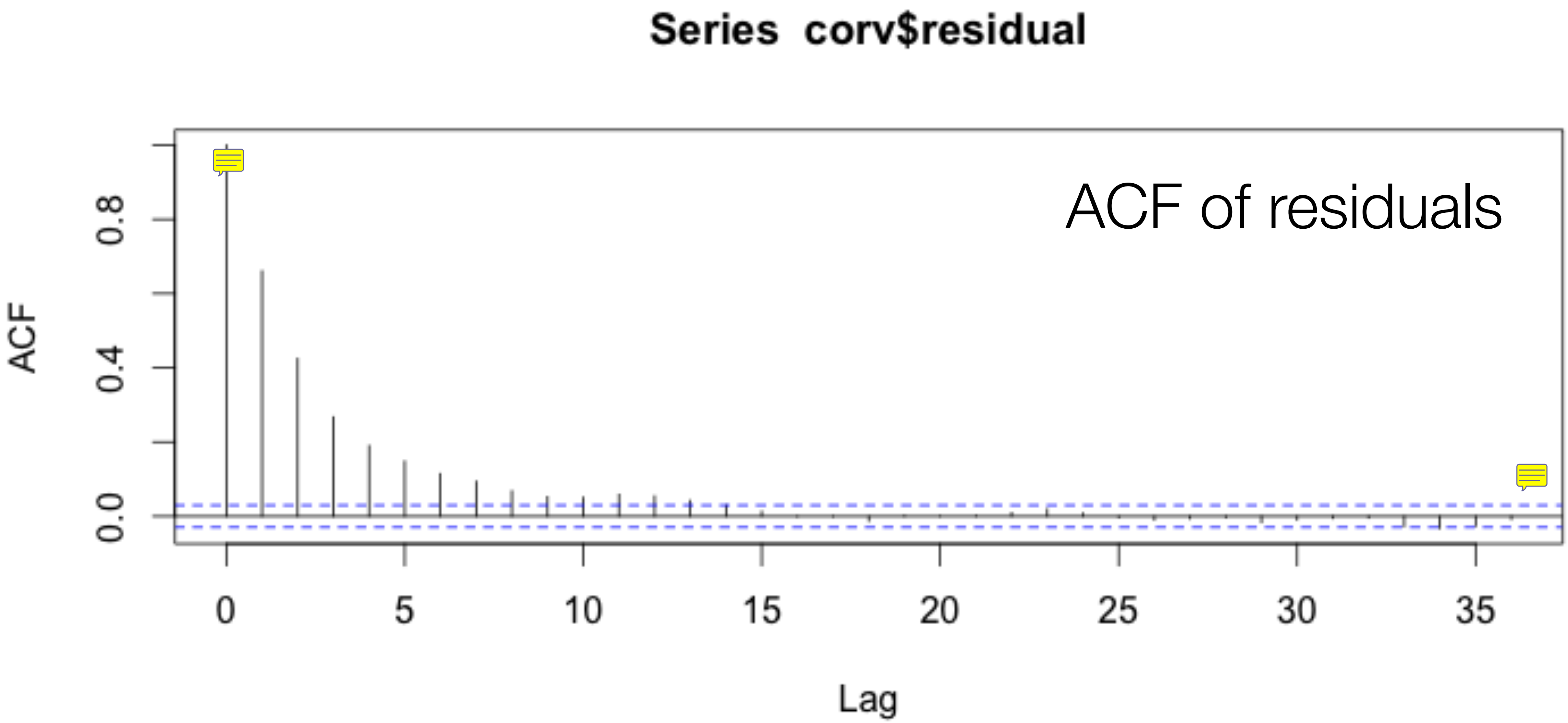
One is AR(1), $\alpha_1 = 0.6$

The other is ARMA(1, 1), $\beta_1 = 0.5$, $\alpha_1 = 0.5$



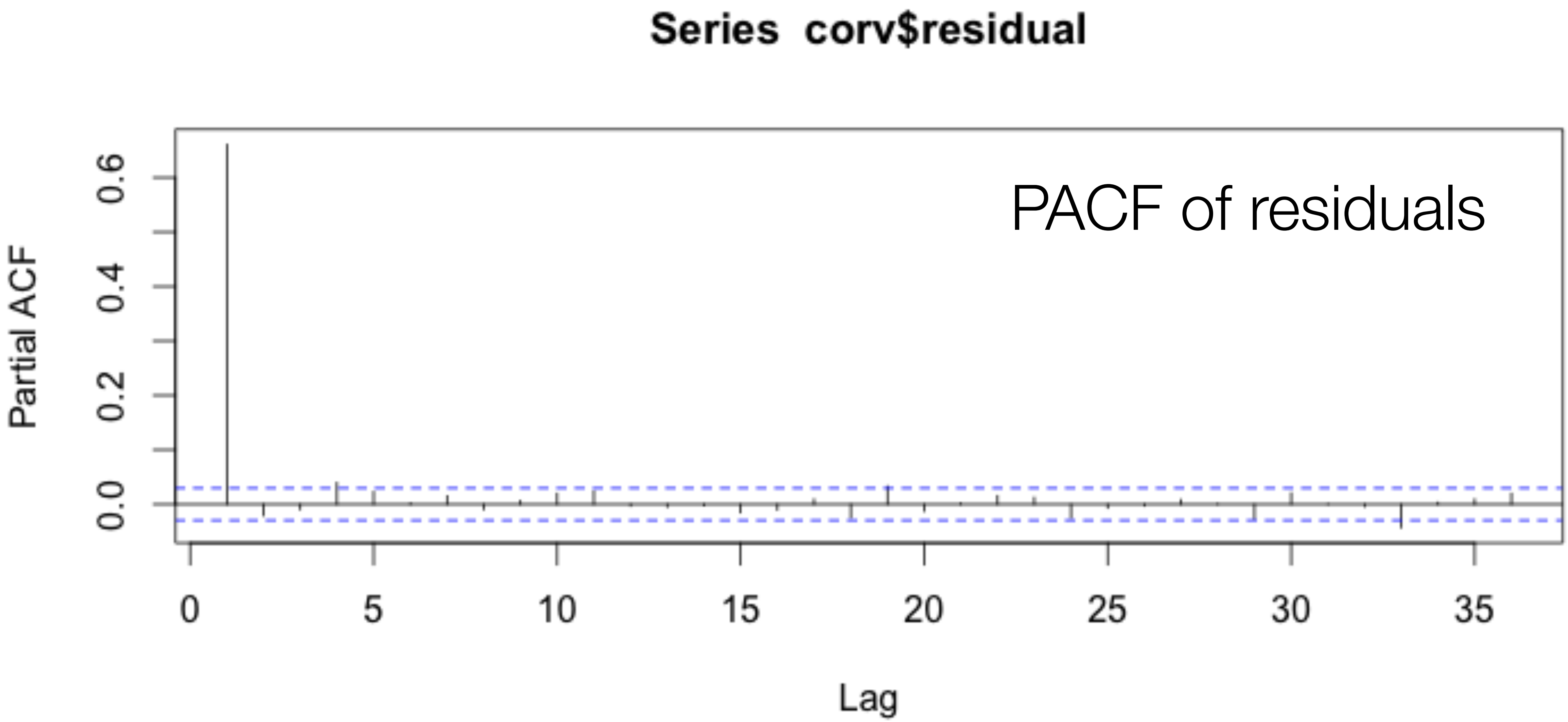
Which is which?

Corvallis temperature



Corvallis temperature

```
pacf(corv$residual, na.action = na.pass)
```



AR(1) looks like a good model

How do we fit it? I.e. how do we estimate α_1?

ARMA(p,q)

Assume we know the order of our process, i.e. we know p and q.

How do we estimate the βs, αs, and σ^2?

Your turn

Name three common approaches to finding an estimate:

Hints:

M ethod of M oment

L east S quare

M aximum L ileklhood

The default in R
Has nice properties

M_______ ___ M___________

We have the theoretical ACF in terms of α, β and σ.

Set the theoretical ACF equal to our sample ACF and solve for the parameters.

So, we don't get too confused, let $\rho(h)$ denote the theoretical ACF, and $r(h)$ the sample ACF.

Corvallis temperature

Assume the residuals can be modelled by an AR(1).

$\rho(1) = \phi$ $\qquad\phi = \alpha$, notation slip

$r(1) = \hat{\phi} = 0.6607$

AR(p)

For AR(p) processes we write down a recursion,

$$\rho(h) = \phi_1 \rho(h-1) + \ldots + \phi_p \rho(h-p), \quad h = 1, \ldots, p$$

$$\sigma^2 = \gamma(0) \left(1 - \phi_1 \rho(1) - \ldots - \phi_p \rho(p)\right)$$

The Yule-Walker equations

Derive Yule-Walker eqns

Yule-Walker Estimates

$$\begin{aligned}
r(1) &= \hat{\phi}_1 + \hat{\phi}_2 r(1) + \ldots + \hat{\phi}_p r(p) \\
r(2) &= \hat{\phi}_1 r(1) + \hat{\phi}_2 + \ldots + \hat{\phi}_p r(p-2) \\
&\vdots \\
r(p) &= \hat{\phi}_1 r(p-1) + \hat{\phi}_2 r(p-2) + \ldots + \hat{\phi}_p
\end{aligned}$$

A set of p equations in p unknowns, solve for $\hat{\phi}_1$ to $\hat{\phi}_p$.

MA(q) and ARMA(p,q)

The method of moments approach gets complicated. End up with non-linear equations to be solved numerically.

The method of moments estimators have bad properties for MA and ARMA processes anyway, so we'll leave it here.

Your turn

Remember linear regression?

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

We can find estimates for β_0 ,β_1 by minimizing

$$\sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2$$

L________ S__________

Consider the AR(1) process

$x_t = \alpha_1 x_{t-1} + w_t$

Define the residual, $e_t = x_t - \hat{\alpha}_1 x_{t-1}$

We could consider finding the $\hat{\alpha}_1$ that minimises the sum of squared residuals,

$$\sum_{t=2}^{n} (x_t - \hat{\alpha}_1 x_{t-1})^2 = \sum_{t=2}^{n} e_t^2$$

called conditional least squares

since we don't see x_0

L_______ s_______ for MA

In general we can always define these residuals, but for MA and ARMA processes they are recursive. For example, MA(1)

$$e_1 = x_1 + \beta_1 e_0$$ assume $e_0 = 0$

$$e_2 = x_2 + \beta_1 e_1$$

$$\vdots$$

$$e_n = x_n + \beta_1 e_{n-1}$$

L_______ s_________ in ARMA

For a general ARMA(p,q):

$$e_t = x_t - \alpha_1 x_{t-1} - \ldots - \alpha_p x_{t-p}$$
$$+ \beta_1 e_{t-1} + \ldots + \beta_q e_{t-q}$$

And we have to set $e_p = \ldots = e_{p+1-q} = 0$, and sum starting at $t = p$, to avoid the x_t we haven't observed

The minimization is done numerically

M__________ L__________

Assume a distribution for the white noise (usually Gaussian), then write the joint density function of our data as a function of the parameters, the likelihood,

$$L(\beta, \theta, \sigma^2) = f(x_1, x_2, \ldots x_n; \beta, \theta, \sigma^2)$$

Find the parameters that maximise the likelihood.

For the non-statisticans: The joint density, f, tells us the probability of our data given certain parameter values. The likelihood, L, tells us how *likely* certain parameters are given our data (f and L are the same function, we just switch what we consider to be the variable). We estimate the parameters by choosing the most likely parameters given the data we saw.

M__________ L___________

Assuming our white noise is Gaussian then the ARMA(p, q), x_t, process is also Gaussian and the likelihood is

$$x_t \sim Normal_n(\mathbf{0}, \Sigma)$$

where $$\Sigma_{ij} = Cov(x_i, x_j) = \gamma(|i-j|)$$

It's complicated, but there are general algorithms for maximizing it.

Maximum Likelihood

The way the function arima in R does it by default.

Nice asymptotic properties, deals with missing data easily.

Always lower variance than method of moments.

Fitting in R

```
arima(ts, order = c(p, d, q))
```

ignore for now

Thursday

I’m out of town.

Chris will lead lecture.

Bring laptops!