# STA1341

# Metode Peramalan Deret Waktu

## Minggu 7
## Model Diagnostics and Forecasting

**Prodi Statistika dan Sains Data**

**Sekolah Sains Data, Matematika, dan Informatika**

**IPB University**

IPB University
Bogor Indonesia

# Outline

- ❑ **Model Diagnostics**
  - ❑ **Residual Analysis**
    - ❑ Plots of the Residuals
    - ❑ Normality of the Residuals
    - ❑ Autocorrelation of the Residuals
    - ❑ The Ljung-Box test
  - ❑ **Overfitting**
- ❑ **Forecasting**

# Diagnostic Model: Residual Analysis

- An AR(2) model with a constant:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \theta_0 + e_t$$

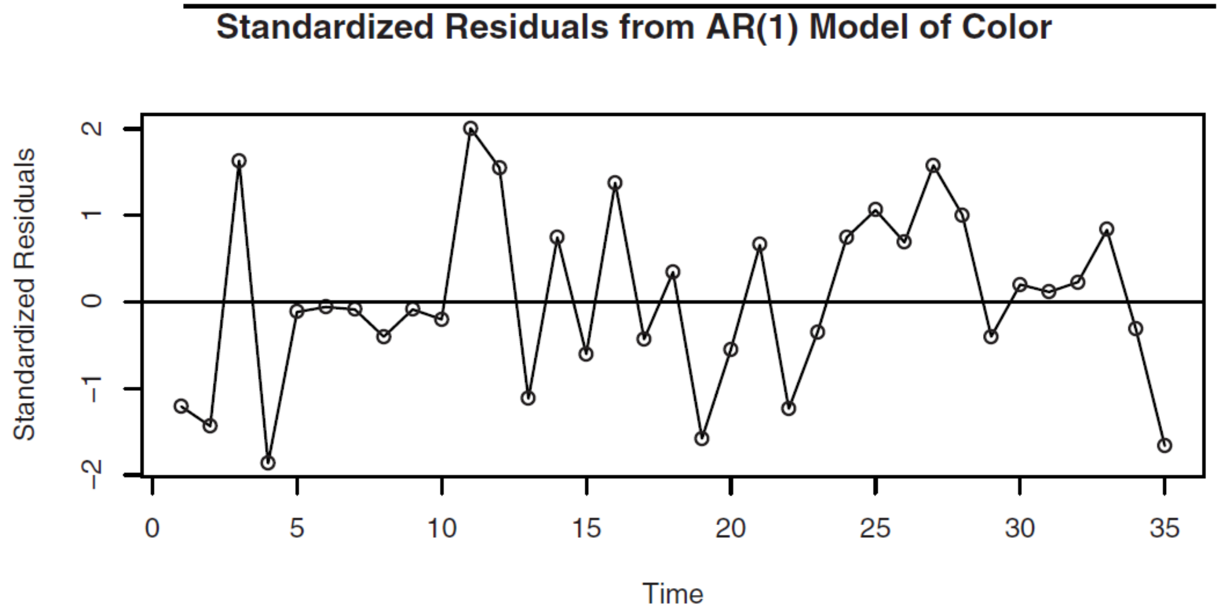
- Having estimated $\phi_1, \phi_2,$ and $\theta_0$, the residual are defined as

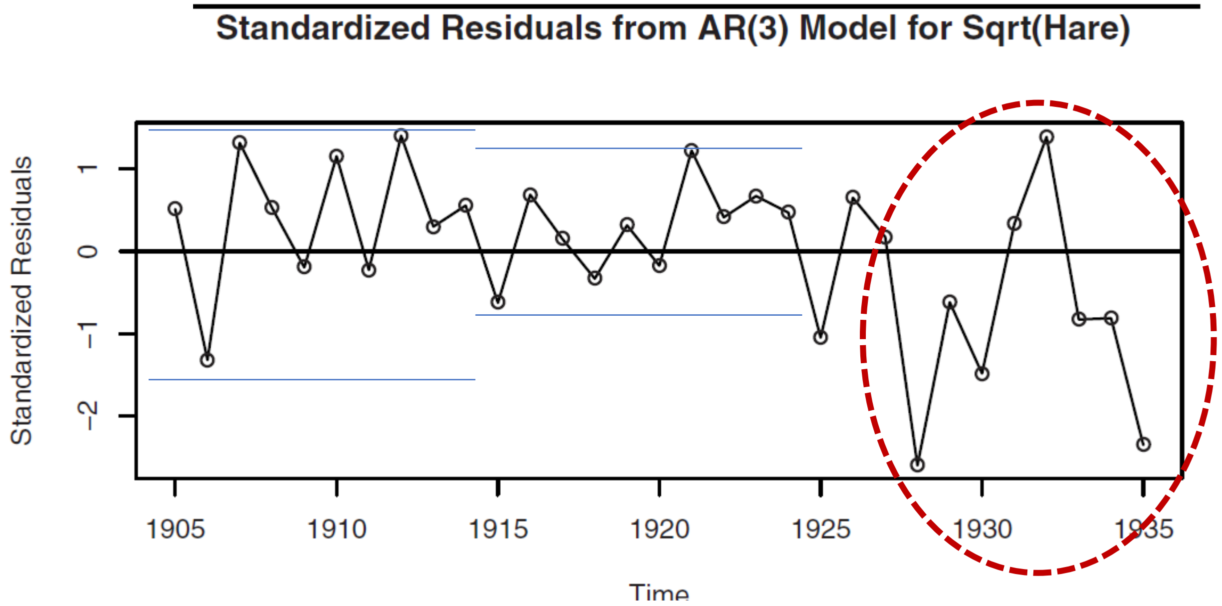
$$\hat{e}_t = Y_t - \hat{\phi}_1 Y_{t-1} - \hat{\phi}_2 Y_{t-2} - \hat{\theta}_0$$

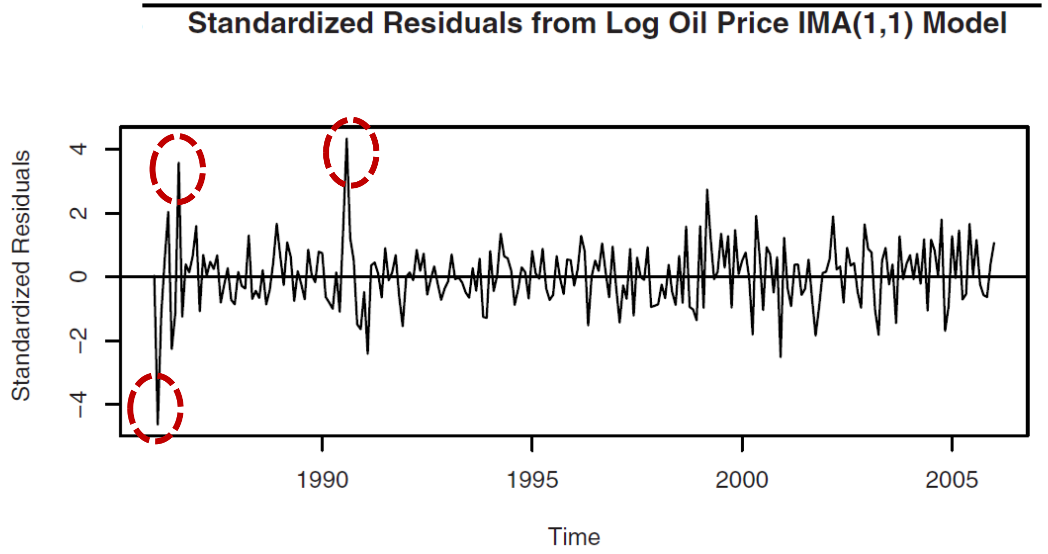
# Diagnostic Model: Residual Analysis,

- Residual = actual – predicted
- If the model is correctly specified and the parameter estimates are reasonably close to the true values, then the residuals should have nearly the properties of white noise
- They should behave roughly like independent, identically distributed normal variables with zero means and common standard deviations

# Model Diagnostics:
# Residual Analysis : Plots of the Residuals



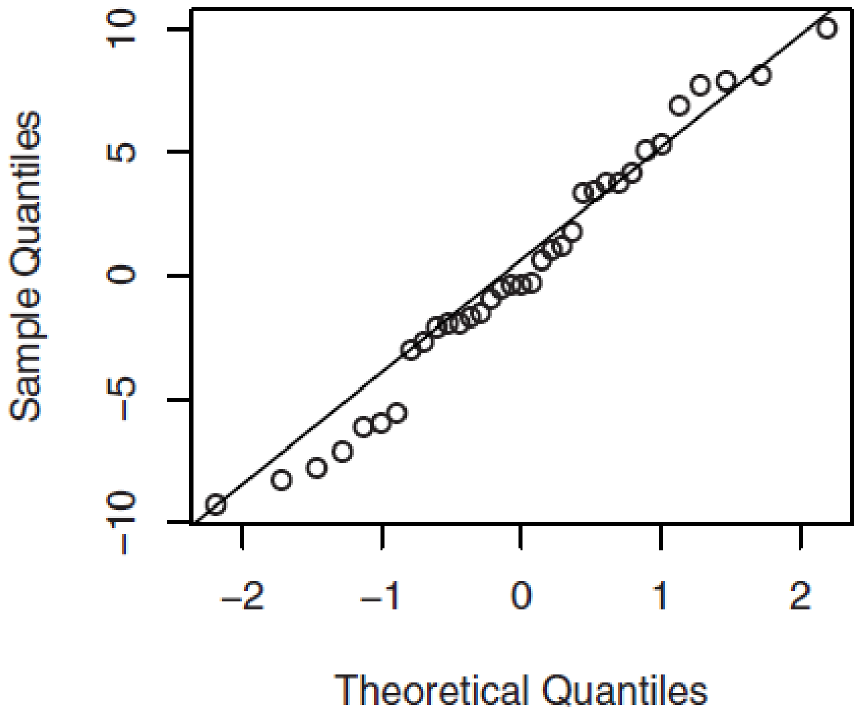

If the model is adequate, we expect the plot to suggest a rectangular scatter around a zero horizontal level with no trends whatsoever.

Cryer and Chan (2008), page 176-178

# Model Diagnostics:
# Residual Analysis : Normality of the Residuals

- The quantile-quantile plots are an effective tool for assessing normality
- The points seem to follow the straight line fairly closely—especially the extreme values.
- This graph would not lead us to reject normality of the error terms in this model.
- The Shapiro-Wilk normality test applied to the residuals produces a test statistic of $W$ = 0.9754, which corresponds to a $p$-value of 0.6057, and we would not reject normality based on this test.

**Exhibit 8.4 Quantile-Quantile Plot: Residuals from AR(1) Color Model**




The Shapiro-Wilk normality test

$H_0$ : states that the variable is normally distributed,

$H_1$ :states that the variable is NOT normally distributed.

# Model Diagnostics:
# Residual Analysis : Normality of the Residuals

**Exhibit 8.5** **Quantile-Quantile Plot: Residuals from AR(3) for Ha**

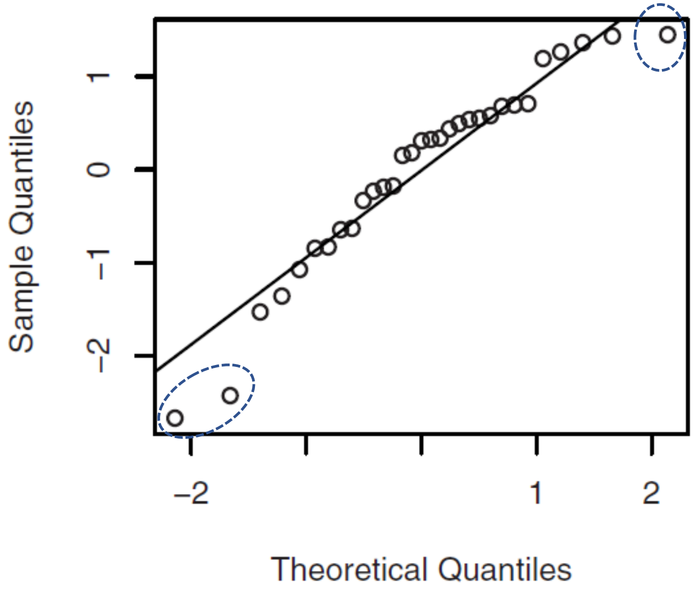
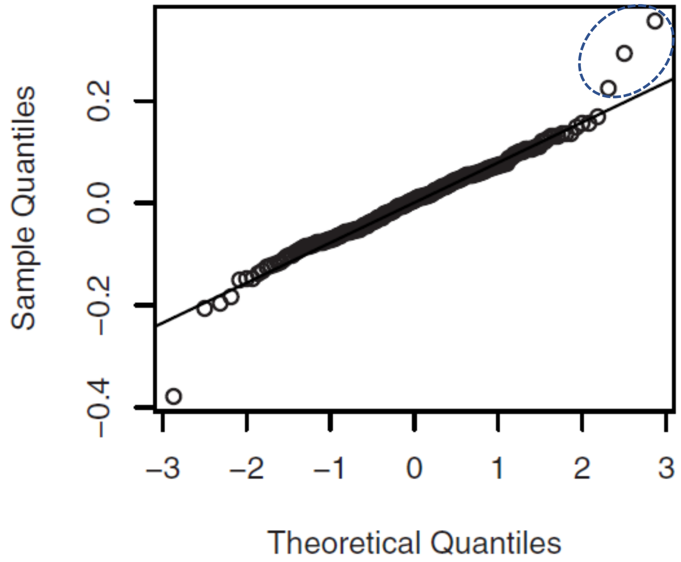



**Exhibit 8.6** **Quantile-Quantile Plot: Residuals from IMA(1,1) Model for Oil**



# Model Diagnostics:
# Residual Analysis : Autocorrelation of the Residuals

- To check on the independence of the noise terms in the model, we consider the sample autocorrelation function of the residuals, denoted $\hat{r}_k$.
- From $Var(r_k) \approx \frac{1}{n}$ and $Corr(r_k, r_j) \approx 0$ for $k \neq j$ we know that for true white noise and large *n*, the sample autocorrelations are approximately uncorrelated and normally distributed with zero means and variance $1/n$.
- As an example, consider a correctly specified and efficiently estimated AR(1) model. It can be shown that, for large *n*,

$$Var(\hat{r}_1) \approx \frac{\phi^2}{n}$$

$$Var(\hat{r}_k) \approx \frac{1-(1-\phi^2)\phi^{2k-2}}{n} \quad \text{for } k > 1$$

$$Corr(\hat{r}_1, \hat{r}_k) \approx -sign(\phi)\frac{(1-\phi^2)\phi^{k-2}}{1-(1-\phi^2)\phi^{2k-2}} \quad \text{for } k > 1$$

$$sign(\phi) = \begin{cases} 1 & \text{if } \phi > 0 \\ 0 & \text{if } \phi = 0 \\ -1 & \text{if } \phi < 0 \end{cases}$$

# Model Diagnostics:
# Residual Analysis : Autocorrelation of the Residuals



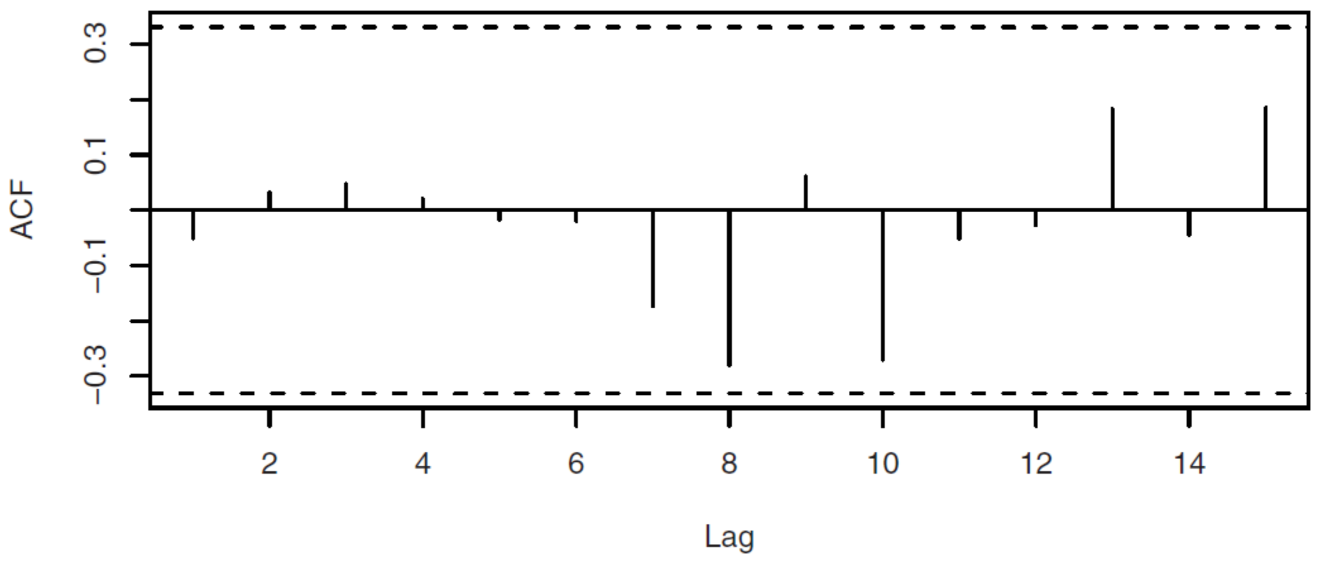
The industrial color property time series with = 0.57 and $n$ = 35

**Exhibit 8.8 Approximate Standard Deviations of Residual ACF values**

| Lag $k$ | 1 | 2 | 3 | 4 | 5 | > 5 |
|---|---|---|---|---|---|---|
| $\sqrt{V\hat{a}r(\hat{r}_k)}$ | 0.096 | 0.149 | 0.163 | 0.167 | 0.168 | 0.169 |

There is no evidence of autocorrelation in the residuals of this model.

**Exhibit 8.9 Sample ACF of Residuals from AR(1) Model for Color**



# Model Diagnostics: Residual Analysis : The Ljung-Box Test

H0 : the error terms are uncorrelated
H1 : the error terms are correlated

$$Q = n(\hat{r}_1^2 + \hat{r}_2^2 + \cdots + \hat{r}_K^2)$$

Box-Pierce Test

$$Q_* = n(n+2)\left(\frac{\hat{r}_1^2}{n-1} + \frac{\hat{r}_2^2}{n-2} + \cdots + \frac{\hat{r}_K^2}{n-K}\right)$$

Ljung-Box Test

$Q$ *or* $Q_*$ has an approximate chi-square distribution with $K - p - q$

# Model Diagnostics:
# Residual Analysis : The Ljung-Box Test

**Exhibit 8.11 Residual Autocorrelation Values from AR(1) Model for Color**

| Lag $k$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Residual ACF | –0.051 | 0.032 | 0.047 | 0.021 | –0.017 | –0.019 |

```
> acf(residuals(m1.color),plot=F)$acf
> signif(acf(residuals(m1.color),plot=F)$acf[1:6],2)
> # display the first 6 acf values to 2 significant digits
```

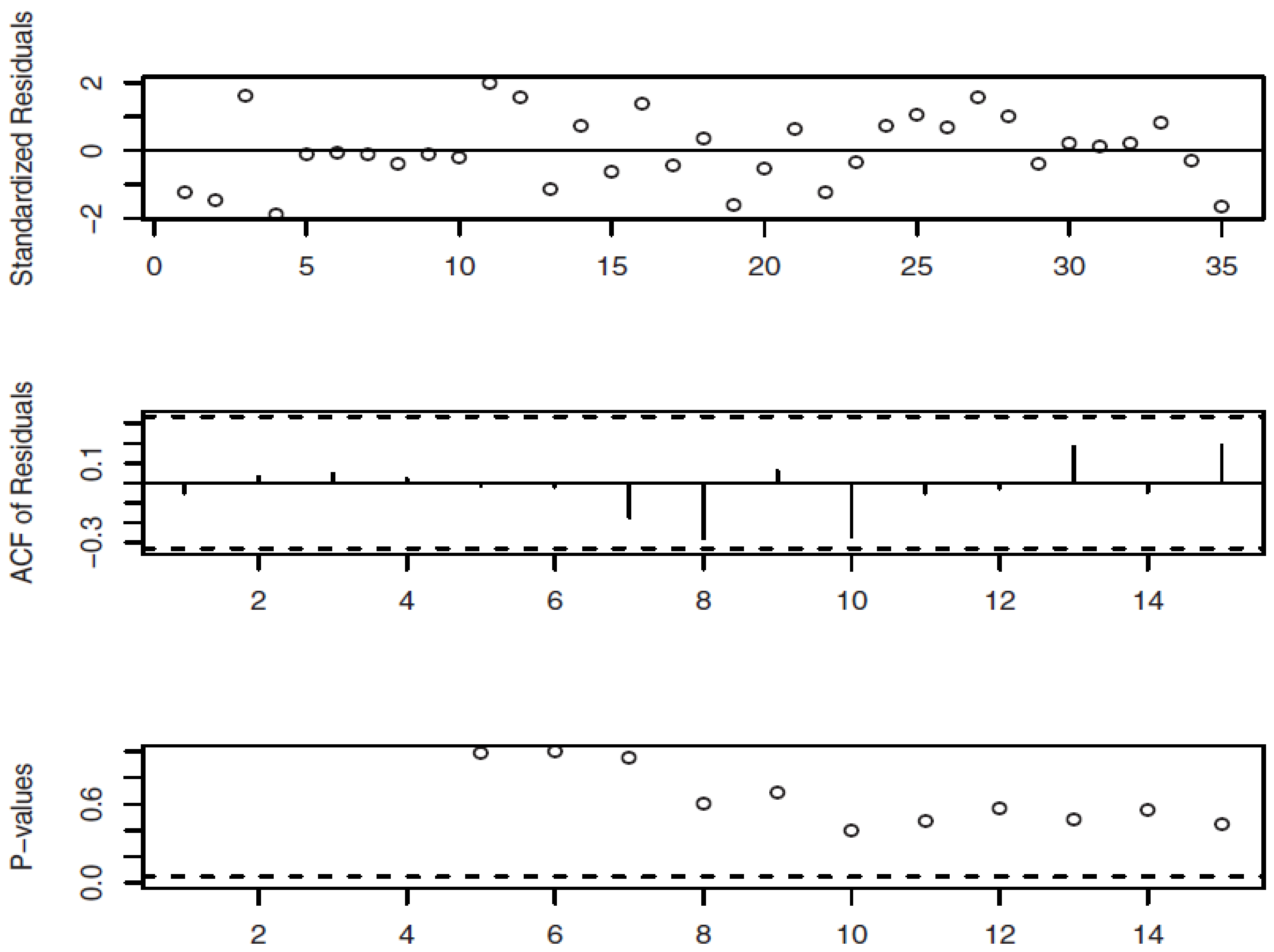
H0 : the error terms are uncorrelated
H1 : the error terms are correlated

The Ljung-Box test statistic with $K = 6$ is equal to

$$Q_* = 35(35+2)\left(\frac{(-0.051)^2}{35-1} + \frac{(0.032)^2}{35-2} + \frac{(0.047)^2}{35-3} + \frac{(0.021)^2}{35-4} + \frac{(-0.017)^2}{35-5} + \frac{(-0.019)^2}{35-6}\right) \approx 0.28$$

This is referred to a chi-square distribution with 6 − 1 = 5 degrees of freedom.
The $p$-value= 0.998

# Exhibit 8.12 Diagnostic Display for the AR(1) Model of Color Property




- A plot of the standardized residuals,
- The sample ACF of the residuals, and
- *p*-values for the Ljung-Box test statistic for a whole range of values of *K* from 5 to 15.
- The horizontal dashed line at 5% helps judge the size of the *p*-values.
- In this instance, everything looks very good.
- The estimated AR(1) model seems to be capturing the dependence structure of the color property time series quite well.

# Model Diagnostics: Overfitting

- After specifying and fitting what we believe to be an adequate model, we fit a slightly more general model; that is, a model "close by" that contains the original model as a special case.
- For example, if an AR(2) model seems appropriate, we might over fit with an AR(3) model. The original AR(2) model would be confirmed if:
  - The estimate of the additional parameter, $\phi_3$, is not significantly different from zero, and
  - The estimates for the parameters in common, $\phi_1$ and $\phi_2$, do not change significantly from their original estimates.

# Model Diagnostics: Overfitting

**13 AR(1) Model Results for the Color Property Series**

| Coefficients:[†] | ar1 | Intercept[‡] |
|---|---|---|
| | 0.5705 | 74.3293 |
| s.e. | 0.1435 | 1.9151 |

sigma^2 estimated as 24.83: log-likelihood = -106.07, AIC = 216.15

**Exhibit 8.14 AR(2) Model Results for the Color Property Series**

| Coefficients: | ar1 | ar2 | Intercept |
|---|---|---|---|
| | 0.5173 | 0.1005 | 74.1551 |
| s.e. | 0.1717 | 0.1815 | 2.1463 |

sigma^2 estimated as 24.6: log-likelihood = -105.92, AIC = 217.84

- ❑ The estimate of $\phi_2$ is not statistically different from zero -- This fact supports the choice of the AR(1) Model
- ❑ We note that the two estimates of $\phi_1$ are quite close
- ❑ The AR(1) fit has a smaller AIC value

# Model Diagnostics: Overfitting

IPB University — Bogor Indonesia —

**AR(1) Model Results for the Color Property Series**

| Coefficients:† | ar1 | Intercept‡ |
|---|---|---|
| | 0.5705 | 74.3293 |
| s.e. | 0.1435 | 1.9151 |

sigma^2 estimated as 24.83: log-likelihood = -106.07, AIC = 216.15

**Overfit of an ARMA(1,1) Model for the Color Series**

| Coefficients: | ar1 | ma1 | Intercept |
|---|---|---|---|
| | 0.6721 | –0.1467 | 74.1730 |
| s.e. | 0.2147 | 0.2742 | 2.1357 |

sigma^2 estimated as 24.63: log-likelihood = -105.94, AIC = 219.88

# Model Diagnostics: Overfitting

IPB University
Bogor Indonesia

**AR(1) Model Results for the Color Property Series**

| Coefficients:[†] | ar1 | Intercept[‡] |
|---|---|---|
| | 0.5705 | 74.3293 |
| s.e. | 0.1435 | 1.9151 |

sigma^2 estimated as 24.83: log-likelihood = -106.07, AIC = 216.15

**Overfitted ARMA(2,1) Model for the Color Property Series**

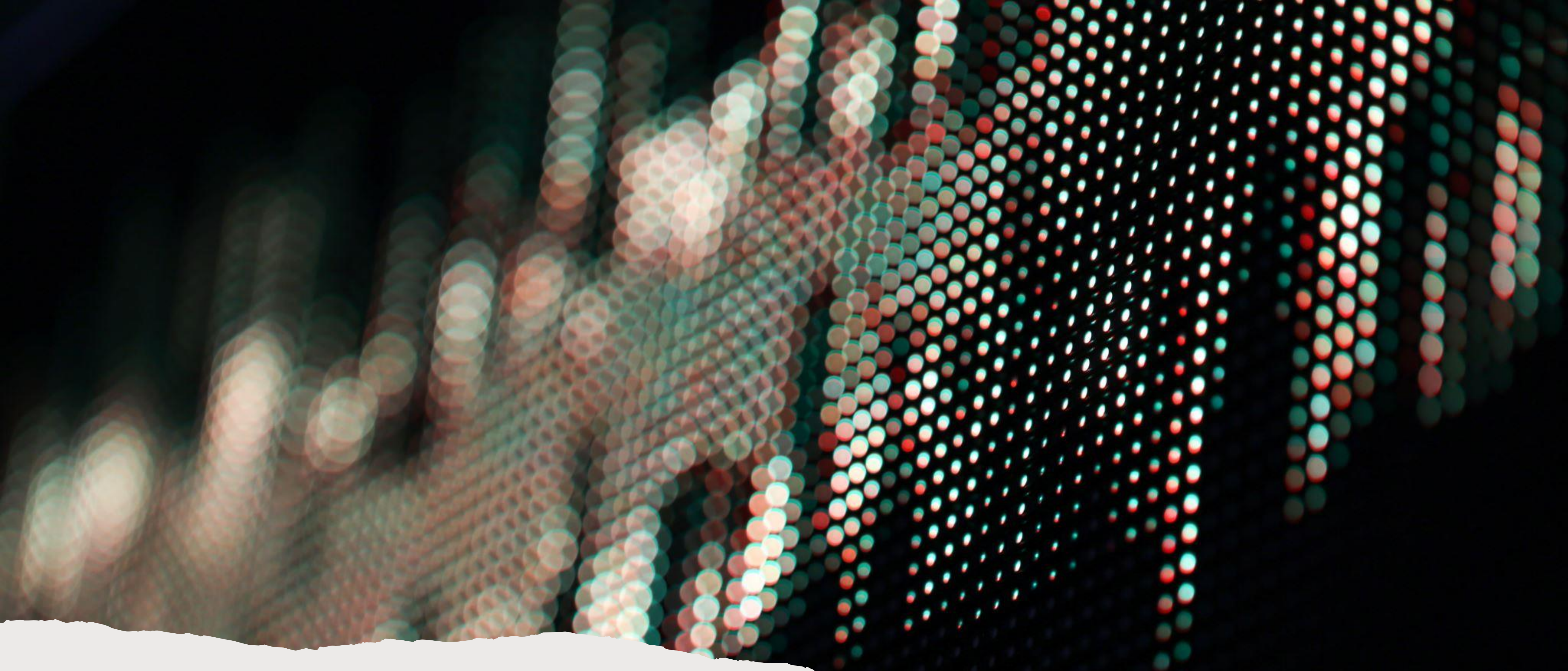
| Coefficients: | ar1 | ar2 | ma1 | Intercept |
|---|---|---|---|---|
| | 0.2189 | 0.2735 | 0.3036 | 74.1653 |
| s.e. | 2.0056 | 1.1376 | 2.0650 | 2.1121 |

sigma^2 estimated as 24.58: log-likelihood = −105.91, AIC = 219.82

# Model Diagnostics: Overfitting

The implications for fitting and overfitting models are as follows:

- Specify the original model carefully. If a simple model seems at all promising, check it out before trying a more complicated model.
- When overfitting, do not increase the orders of both the AR and MA parts of the model simultaneously.
- Extend the model in directions suggested by the analysis of the residuals. For example, if after fitting an MA(1) model, substantial correlation remains at lag 2 in the residuals, try an MA(2), not an ARMA(1,1).



# Forecasting

# Forecasting – AR(1)

- One of the primary objectives of building a model for a time series is to be able to forecast the values for that series at future times.
- Of equal importance is the assessment of the precision of those forecasts.
- Based on the available history of the series up to time $t$, namely $Y_1, Y_2, \dots, Y_{t-1}, Y_t$, we would like to forecast the value of $Y_{t+l}$ that will occur $l$ time units into the future.
- We call time $t$ the **forecast origin** and $l$ the **lead time** for the forecast, and denote the forecast itself as $\hat{Y}_t(l)$.
- The minimum mean square error forecast is given by

$$\hat{Y}_t(\ell) = E(Y_{t+\ell} | Y_1, Y_2, \dots, Y_t)$$

# Forecasting – AR(1)

We shall first illustrate many of the ideas with the simple AR(1) process with a nonzero mean that satisfies

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + e_t \tag{9.3.1}$$

Consider the problem of forecasting one time unit into the future. Replacing $t$ by $t + 1$ in Equation (9.3.1), we have

$$Y_{t+1} - \mu = \phi(Y_t - \mu) + e_{t+1} \tag{9.3.2}$$

Given $Y_1, Y_2, \ldots, Y_{t-1}, Y_t$, we take the conditional expectations of both sides of Equation (9.3.2) and obtain

$$\hat{Y}_t(1) - \mu = \phi[E(Y_t|Y_1, Y_2, \ldots, Y_t) - \mu] + E(e_{t+1}|Y_1, Y_2, \ldots, Y_t) \tag{9.3.3}$$

# Forecastingn – AR(1)

$$\hat{Y}_t(1) - \mu = \phi[E(Y_t|Y_1, Y_2, \ldots, Y_t) - \mu] + E(e_{t+1}|Y_1, Y_2, \ldots, Y_t) \qquad (9.3.3)$$

Now, from the properties of conditional expectation, we have

$$E(Y_t|Y_1, Y_2, \ldots, Y_t) = Y_t \qquad (9.3.4)$$

Also, since $e_{t+1}$ is independent of $Y_1, Y_2, \ldots, Y_{t-1}, Y_t$, we obtain

$$E(e_{t+1}|Y_1, Y_2, \ldots, Y_t) = E(e_{t+1}) = 0 \qquad (9.3.5)$$

Thus, Equation (9.3.3) can be written as

$$\hat{Y}_t(1) = \mu + \phi(Y_t - \mu) \qquad (9.3.6)$$

# Forecasting – AR(1)

$$Y_{t+l} - \mu = \phi(Y_{t+l-1} - \mu) + e_{t+l}$$

$l = 2$

$$Y_{t+2} - \mu = \phi(Y_{t+1} - \mu) + e_{t+2}$$

$$\hat{Y}_t(2) - \mu = \phi\left(\hat{Y}_t(1) - \mu\right) + 0$$
$$\hat{Y}_t(2) = \phi\left(\hat{Y}_t(1) - \mu\right) + \mu$$

$$\hat{Y}_t(3) = \phi\left(\hat{Y}_t(2) - \mu\right) + \mu$$
$$\hat{Y}_t(3) = \phi\left(\phi\left(\hat{Y}_t(1) - \mu\right) + \mu \ - \mu\right) + \mu$$
$$\hat{Y}_t(3) = \phi^2\left(\hat{Y}_t(1) - \mu\right) + \mu$$
$$\hat{Y}_t(3) = \phi^2(\phi(Y_t - \mu) + \mu - \mu) + \mu$$

$$\hat{Y}_t(3) = \phi^3(Y_t - \mu) + \mu$$

# Forecasting – AR(1)

Jniversity
gor Indonesia

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + e_t \tag{9.3.1}$$

Now consider a general lead time $\ell$. Replacing $t$ by $t + \ell$ in Equation (9.3.1) and taking the conditional expectations of both sides produces

$$\hat{Y}_t(\ell) = \mu + \phi[\hat{Y}_t(\ell - 1) - \mu] \quad \text{for } \ell \geq 1 \tag{9.3.7}$$

since $E(Y_{t+\ell-1}|Y_1, Y_2, \ldots, Y_t) = \hat{Y}_t(\ell - 1)$ and, for $\ell \geq 1$, $e_{t+\ell}$ is independent of $Y_1$, $Y_2, \ldots, Y_{t-1}, Y_t$.

$$\begin{aligned}
\hat{Y}_t(\ell) &= \phi[\hat{Y}_t(\ell - 1) - \mu] + \mu \\
&= \phi\{\phi[\hat{Y}_t(\ell - 2) - \mu]\} + \mu \\
&\vdots \\
&= \phi^{\ell-1}[\hat{Y}_t(1) - \mu] + \mu
\end{aligned}$$

or

$$\boxed{\hat{Y}_t(\ell) = \mu + \phi^{\ell}(Y_t - \mu)} \tag{9.3.8}$$

# Forecasting—Example

**Exhibit 9.1 Maximum Likelihood Estimation of an AR(1) Model for Color**

| Coefficients: | ar1 | intercept† |
|---|---|---|
| | 0.5705 | 74.3293 |
| s.e. | 0.1435 | 1.9151 |

sigma^2 estimated as 24.8: log-likelihood = –106.07, AIC = 216.15

†Remember that the intercept here is the estimate of the process mean $\mu$—not $\theta_0$.

The last observed value of the color property is 67, so we would forecast one time period ahead as†

$$\begin{aligned}\hat{Y}_t(1) &= 74.3293 + (0.5705)(67 - 74.3293) \\ &= 74.3293 - 4.181366 \\ &= 70.14793\end{aligned}$$

$\hat{Y}_t(2)$?
$\hat{Y}_t(6)$?
$\hat{Y}_t(8)$?
$\hat{Y}_t(10)$?

In general, since $|\phi| < 1$, we have simply

$$\hat{Y}_t(\ell) \approx \mu \quad \text{for large } \ell$$

# Forecasting - AR(1)

$$Var\big(e_t(l)\big) = \sigma_e^2\left[\frac{1-\phi^{2l}}{1-\phi^2}\right]$$

For long lead times, we have

$Var(e_t(l)) \approx \frac{\sigma_e^2}{1-\phi^2}$ for large $l$ or $Var\big(e_t(l)\big) \approx Var(Y_t) = \gamma_0$ for large $l$

# Forecasting – MA(1)

University
gor Indonesia

$$Y_t = \mu + e_t - \theta e_{t-1}$$

$$Y_{t+1} = \mu + e_{t+1} - \theta e_t$$

$$\hat{Y}_t(1) = E[\mu + e_{t+1} - \theta e_t | \mathrm{Y_t, Y_{t-1}, \dots, Y_1}]$$

$$\hat{Y}_t(1) = \mu + 0 - \theta e_t$$

$$\hat{Y}_t(1) = \mu - \theta e_t$$

$$Y_{t+2} = \mu + e_{t+2} - \theta e_{t+1}$$

$$\hat{Y}_t(2) = E[\mu + e_{t+2} - \theta e_{t+1} | \mathrm{Y_t, Y_{t-1}, \dots, Y_1}]$$

$$\hat{Y}_t(2) = \mu + 0 - 0$$

$$\hat{Y}_t(2) = \mu$$

For longer lead times, we have

$$\hat{Y}_t(\ell) = \mu + E(e_{t+\ell} | Y_1, Y_2, \dots, Y_t) - \theta E(e_{t+\ell-1} | Y_1, Y_2, \dots, Y_t) \Rightarrow \hat{Y}_t(\ell) = \mu \text{ for } \ell > 1$$

# Prediction Limits : ARIMA Models

University
gor Indonesia

- If the white noise terms {*et*} in a general ARIMA series each arise independently from a normal distribution, then the forecast error $e_t(l)$ will also have a normal distribution
- For an AR(1) model, $Var\left(e_t(l)\right) = \sigma_e^2\left[\frac{1-\phi^{2l}}{1-\phi^2}\right]$
- In practice, $\sigma_e^2$ will be unknown and must be estimated from the observed time series.
- As a numerical example, consider the AR(1) model that we estimated for the industrial color property series. We use $\phi = 0.5705, \mu = 74.3293, and\ \sigma_e^2 = 24.8.$

For a one-step-ahead prediction, we have

$$70.14793 \pm 1.96\sqrt{24.8} = 70.14793 \pm 9.760721 \quad \text{or} \quad 60.39 \quad \text{to} \quad 79.91$$

Two steps ahead, we obtain

$$71.86072 \pm 11.88343 \quad \text{or} \quad 60.71 \quad \text{to} \quad 83.18$$

Notice that this prediction interval is wider than the previous interval. Forecasting ten steps ahead leads to

$$74.173934 \pm 11.88451 \quad \text{or} \quad 62.42 \quad \text{to} \quad 86.19$$

By lead 10, both the forecast and the forecast limits have settled down to their long-lead values.

Forecasting Illustrations : An Industrial Chemical Process

Exhibit 1.3 Time Series Plot of Color Property from a Chemical Process

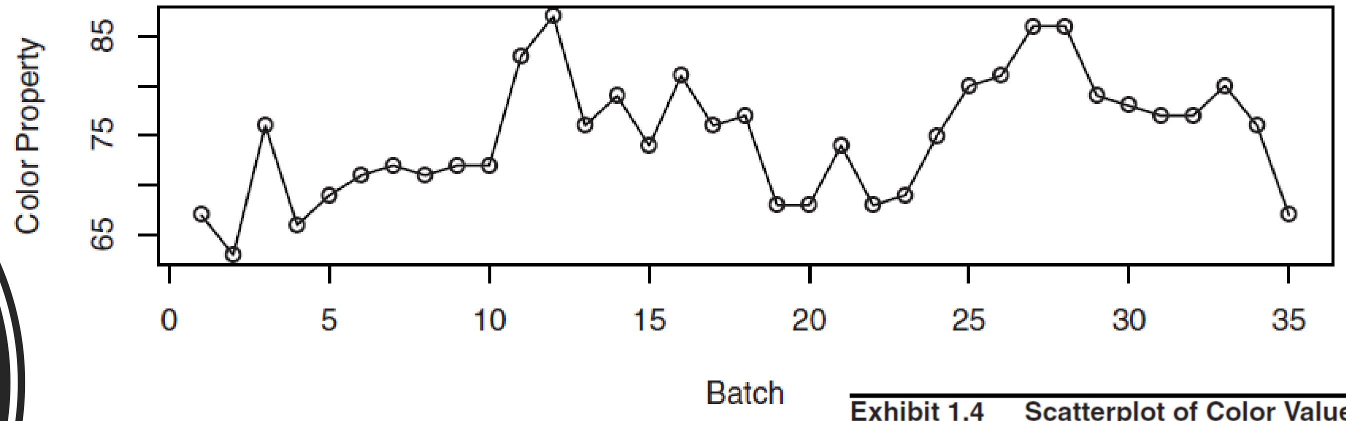
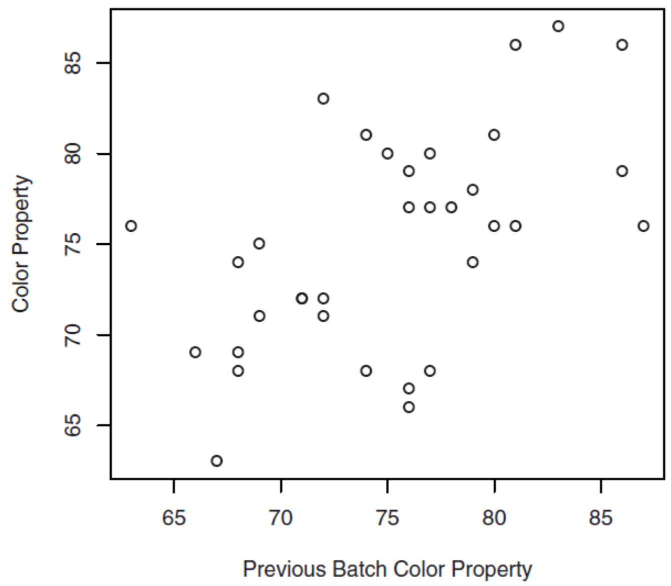



Exhibit 1.4 Scatterplot of Color Value versus Previous Color Value




**Step 1. Stationary**

Exhibit 6.25 Sample ACF for the Color Property Series

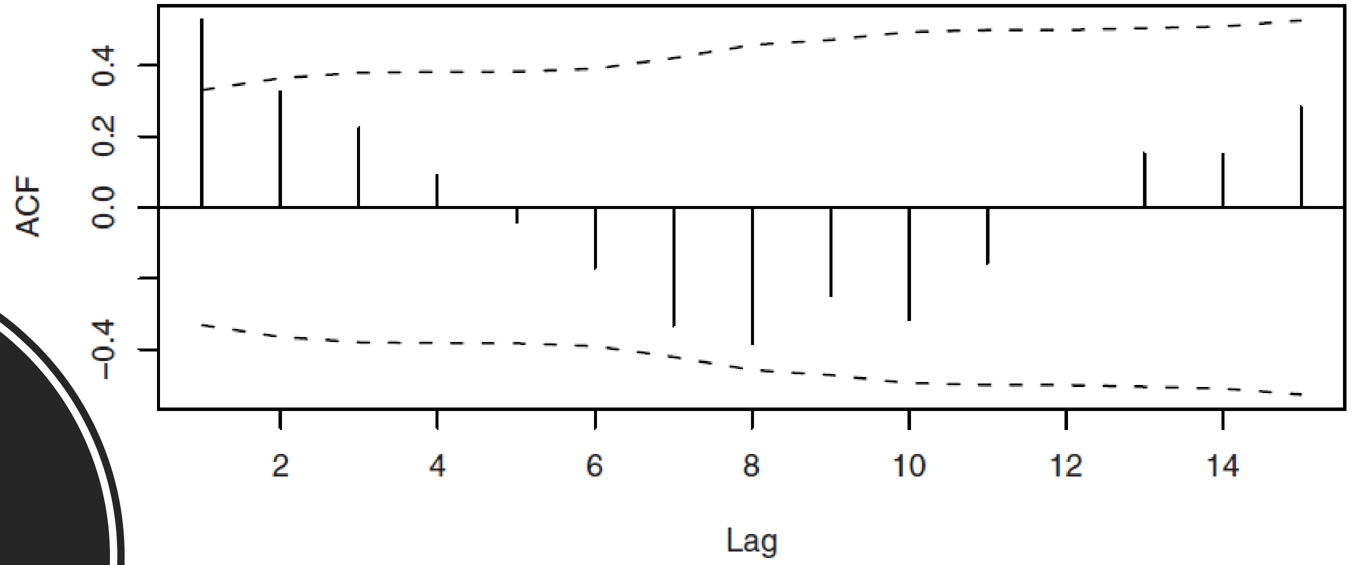
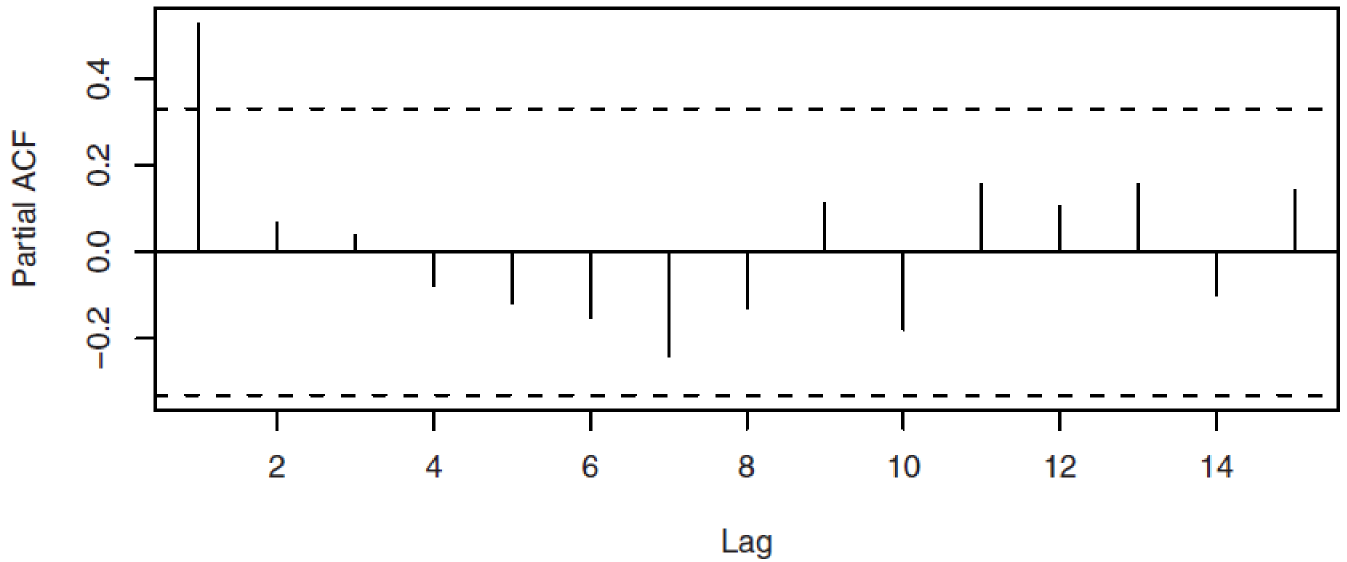



Exhibit 6.26 Sample Partial ACF for the Color Property Series




Forecasting Illustrations : An Industrial Chemical Process

**Step 2. Model Specification**

## Step 3. Parameter Estimation

IPB University — Bogor Indonesia

# Forecasting Illustrations : An Industrial Chemical Process

**Exhibit 7.7 Parameter Estimation for the Color Property Series**

| Parameter | Method-of-Moments Estimate | Conditional SS Estimate | Unconditional SS Estimate | Maximum Likelihood Estimate | *n* |
|---|---|---|---|---|---|
| $\phi$ | 0.5282 | 0.5549 | 0.5890 | 0.5703 | 35 |

## Step 4. Model Diagnostics

**Exhibit 8.1 Standardized Residuals from AR(1) Model of Color**

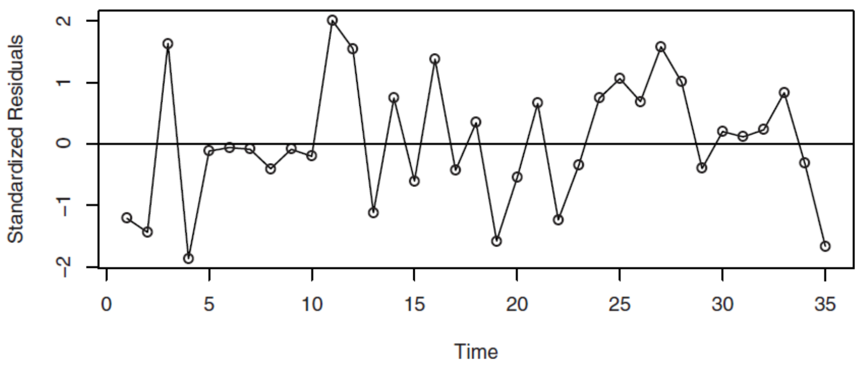



**Exhibit 8.4 Quantile-Quantile Plot: Residuals from AR(1) Color Model**

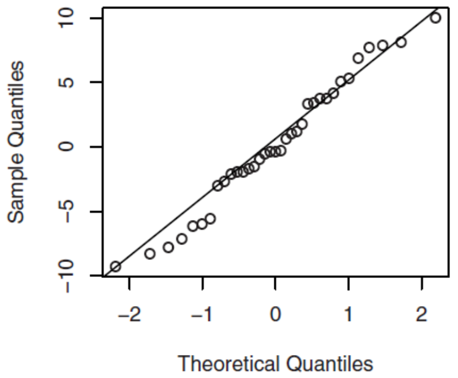
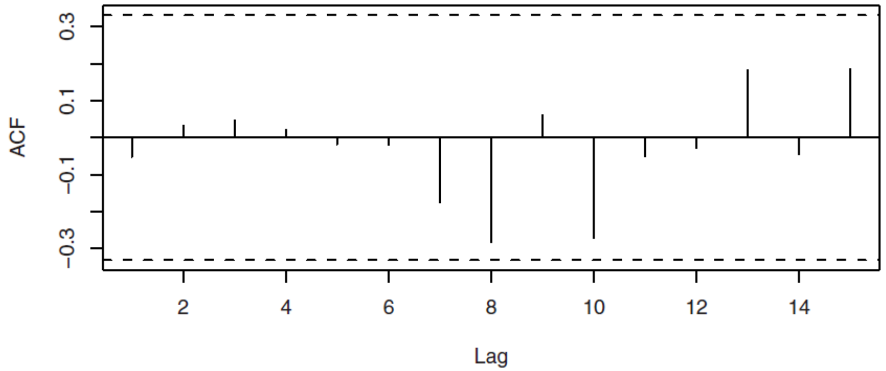



**Exhibit 8.9 Sample ACF of Residuals from AR(1) Model for Color**




The Ljung-Box test statistic with $K = 6$ is equal to

$$Q_* = 35(35+2)\left(\frac{(-0.051)^2}{35-1} + \frac{(0.032)^2}{35-2} + \frac{(0.047)^2}{35-3} + \frac{(0.021)^2}{35-4} + \frac{(-0.017)^2}{35-5} + \frac{(-0.019)^2}{35-6}\right) \approx 0.28$$

This is referred to a chi-square distribution with 6 − 1 = 5 degrees of freedom.
The *p*-value= 0.998

**Step 5.** Overfitting

Forecasting Illustrations : An Industrial Chemical Process

**Exhibit 8.13 AR(1) Model Results for the Color Property Series**

| Coefficients:[†] | ar1 | Intercept[‡] |
|---|---|---|
| | 0.5705 | 74.3293 |
| s.e. | 0.1435 | 1.9151 |

sigma^2 estimated as 24.83: log-likelihood = -106.07, AIC = 216.15

[†] `m1.color` # R code to obtain table

[‡] Recall that the intercept here is the estimate of the process mean $\mu$—not $\theta_0$.

**Exhibit 8.14 AR(2) Model Results for the Color Property Series**

| Coefficients: | ar1 | ar2 | Intercept |
|---|---|---|---|
| | 0.5173 | 0.1005 | 74.1551 |
| s.e. | 0.1717 | 0.1815 | 2.1463 |

sigma^2 estimated as 24.6: log-likelihood = -105.92, AIC = 217.84

**Exhibit 8.15 Overfit of an ARMA(1,1) Model for the Color Series**

| Coefficients: | ar1 | ma1 | Intercept |
|---|---|---|---|
| | 0.6721 | –0.1467 | 74.1730 |
| s.e. | 0.2147 | 0.2742 | 2.1357 |

sigma^2 estimated as 24.63: log-likelihood = -105.94, AIC = 219.88

## **Step 6.** Forecasting

IPB University — Bogor Indonesia

# Forecasting Illustrations : An Industrial Chemical Process

**Exhibit 9.1 Maximum Likelihood Estimation of an AR(1) Model for Color**

| Coefficients: | ar1 | intercept[†] |
|---|---|---|
| | 0.5705 | 74.3293 |
| s.e. | 0.1435 | 1.9151 |

sigma^2 estimated as 24.8: log-likelihood = −106.07, AIC = 216.15

The last observed value of the color property is 67, so we would forecast one time period ahead as[†]

$$\begin{aligned}\hat{Y}_t(1) &= 74.3293 + (0.5705)(67 - 74.3293)\\ &= 74.3293 - 4.181366\\ &= 70.14793\end{aligned}$$

For lead time 2, we have from Equation (9.3.7)

$$\begin{aligned}\hat{Y}_t(2) &= 74.3293 + 0.5705(70.14793 - 74.3293)\\ &= 74.3293 - 2.385472\\ &= 71.94383\end{aligned}$$

Alternatively, we can use Equation (9.3.8):

$$\begin{aligned}\hat{Y}_t(2) &= 74.3293 + (0.5705)^2(67 - 74.3293)\\ &= 71.92823\end{aligned}$$

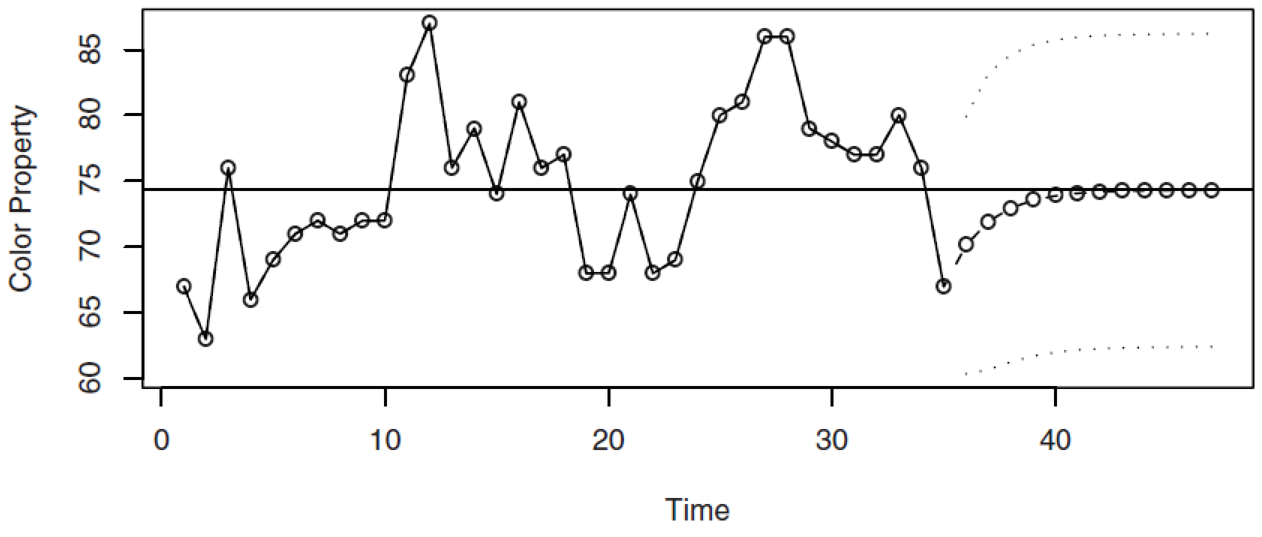
At lead 5, we have

$$\begin{aligned}\hat{Y}_t(5) &= 74.3293 + (0.5705)^5(67 - 74.3293)\\ &= 73.88636\end{aligned}$$

and by lead 10 the forecast is

$$\hat{Y}_t(10) = 74.30253$$

**Exhibit 9.3 Forecasts and Forecast Limits for the AR(1) Model for Color**



# Terima Kasih

IPB University
Bogor Indonesia

INSTITUT PERTANIAN BOGOR

# IPB University

— Bogor Indonesia —

Inspiring Innovation with Integrity
in Agriculture, Ocean and Biosciences for a Sustainable World