Time series analysis

Lecture 2. Autoregressive-moving average models ARMA (p,q)

Dr. Khamidov Obidjon

Autoregressive Integrated Moving Average (ARIMA) Models

1 The Sample Autocorrelation Function
2 Autoregressive Moving Average (ARMA) Models
3 Partial Autocorrelations and Model Selection
4 Model Estimation and Selection
5 Model Diagnostics
6 Outliers
7 Forecasting with ARIMA Models
8 Seasonal ARIMA Models
9 State-Space and ARIMA Models
10 GARCH Models
11 Principles of ARIMA Modeling

The Sample Autocorrelation Function

- The first order (or lag 1) autocorrelation measures the correlation between successive observations in a time series.
- The sample value is computed as:

$$r_1 = \frac{\sum_{t=2}^{n} (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2}$$

Q: Why do we use a common mean?
Q: Why don't we adjust for the fact that the denominator has *n* terms whereas the numerator has only *(n-1)*.

The Sample Autocorrelation Function

Example: Calculation of First-Order Autocorrelation

Week, X	1	2	3	4	5	6	7	Sums
Sales (000s), Y_t	15	10	12	16	9	12	10	84
Lagged values, Y_{t-1}		15	10	12	16	9	12	
$Y_t - \overline{Y}$		−2	0	4	−3	0	−2	
$Y_{t-1} - \overline{Y}$		3	−2	0	4	−3	0	
$(Y_t - \overline{Y})(Y_{t-1} - \overline{Y})$		−6	0	0	−12	0	0	−18
$(Y_t - \overline{Y})^2$	9	4	0	16	9	0	4	42

Total sales over the seven weeks are 84, so the weekly average is 12.

$$r_1 = -18/42 = -0.429$$

The Sample Autocorrelation Function

- Higher order autocorrelations: The sample function for lag *k* is:

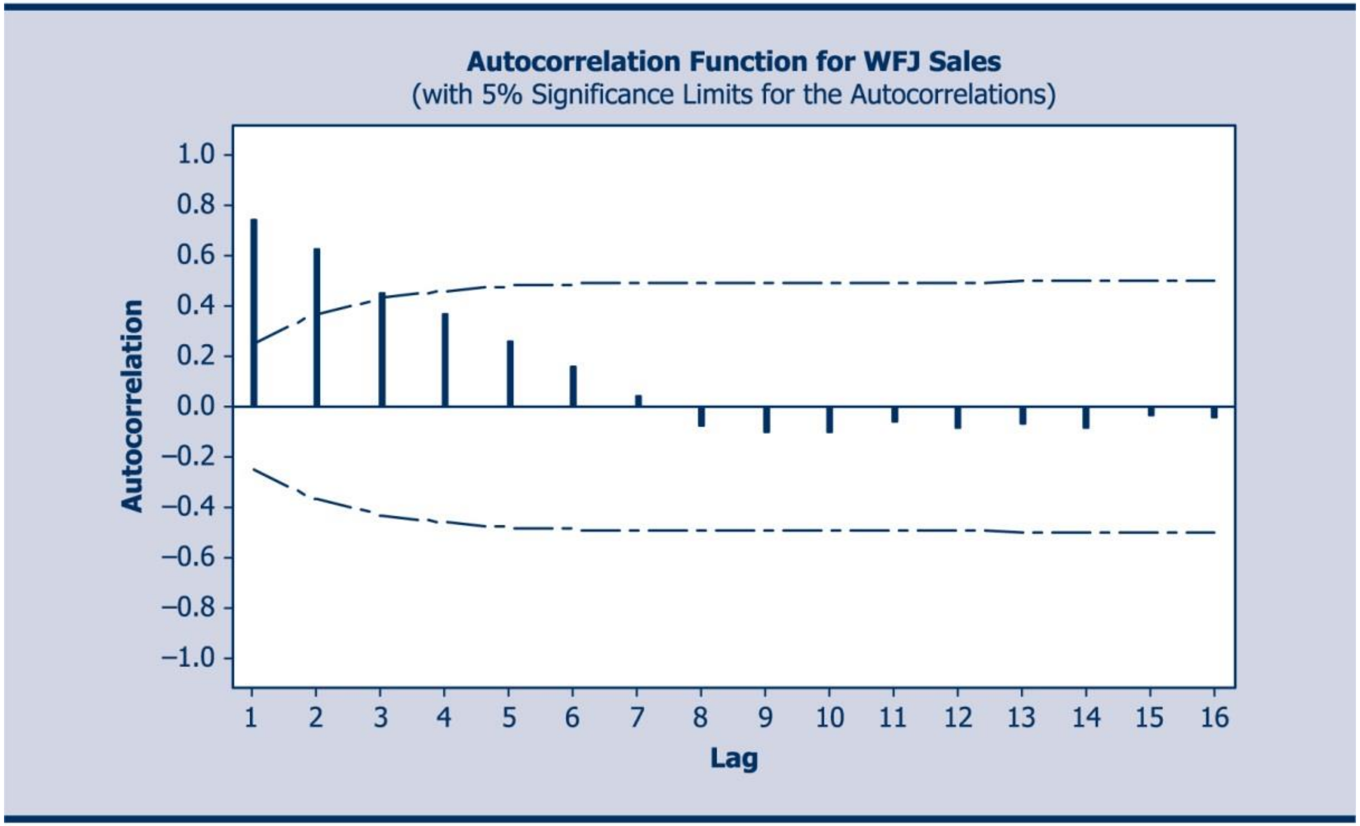
$$r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t-1}^{n} (Y_t - \bar{Y})^2}$$

- The plot of r_k against *k* for k=1,2,… is known as the *sample autocorrelation function (ACF)*, typically plotted for the first *n/4* lags or thereabouts.
- The plot is supplemented with 5% significance limits to enable a graphical check of whether dependence exists at a given lag.

The Sample Autocorrelation Function

Example: The Autocorrelation Function for WFJ Sales

FIGURE 6.1 THE ACF FOR WFJ SALES (MINITAB)



Data shown is from file WFJ_sales.xlsx.

Model Assumptions

- A time series is *stationary* if it has a constant mean and variance and its autocorrelations depend only on the relative time between the observations. Formally, we introduce the notation ω^2 for the variance of Y and ρ_k to denote the autocorrelation at lag k. Then, for stationarity, we require that

$$E(Y_t) = \mu$$

$$V(Y_t) = \omega^2$$

$$Corr(Y_t, Y_{t-k}) = \rho_k$$

Q: Can you suggest series that might be stationary? Also suggest series that may be non-stationary in either the mean or the variance or both.

Autoregressive Moving Average (ARMA) Models

- The autoregressive model of lag 1, written as AR(1) is:

$$Y_t = \delta + \phi Y_{t-1} + \varepsilon_t$$

- Assumptions regarding the error term are the same as before:
  - Zero mean, constant variance, and mutually uncorrelated
  - The model is stationary if and only if

$$-1 < \phi < 1$$

- Stationary models have unconditional mean, variance and autocorrelations that do not depend on time *t*:

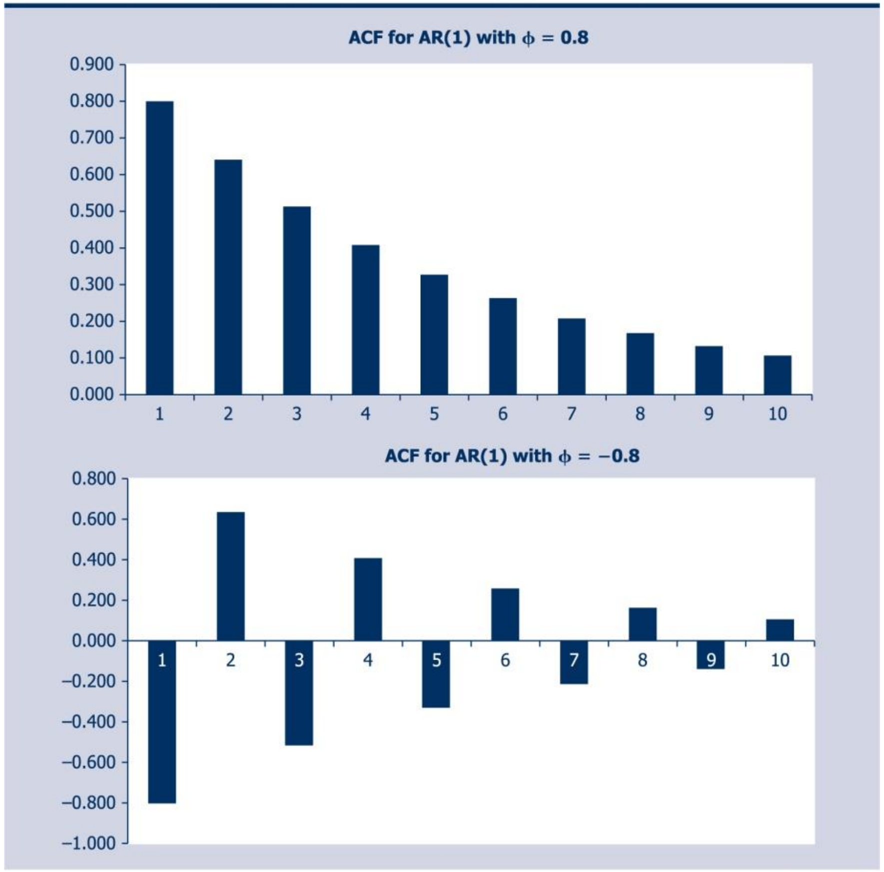
$$E(Y_t) = \mu = \delta / (1 - \phi)$$

$$V(Y_t) = \omega^2 = \sigma^2 / (1 - \phi^2)$$

$$Corr(Y_t, Y_{t-k}) = \phi^k \quad k = 1, 2, \cdots$$

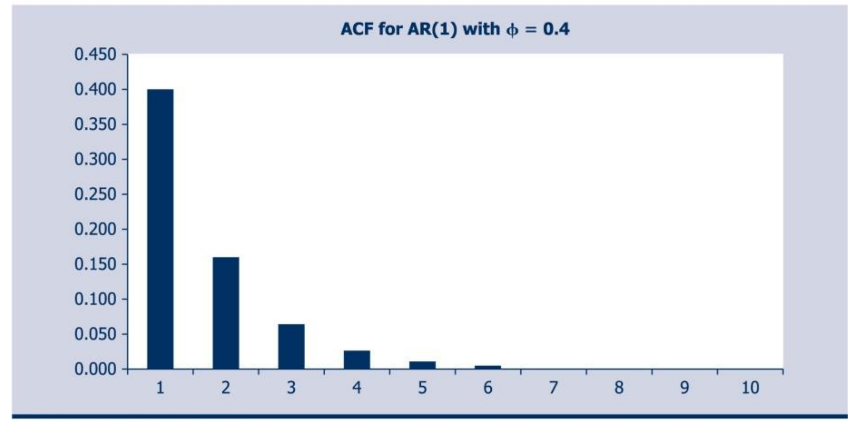
Autoregressive Moving Average (ARMA) Models

FIGURE 6.2 ACF FOR AR(1) SCHEME WITH $\phi = 0.8$, $\phi = -0.8$ AND $\phi = 0.4$



(*Continued*)

FIGURE 6.2 (CONTINUED)



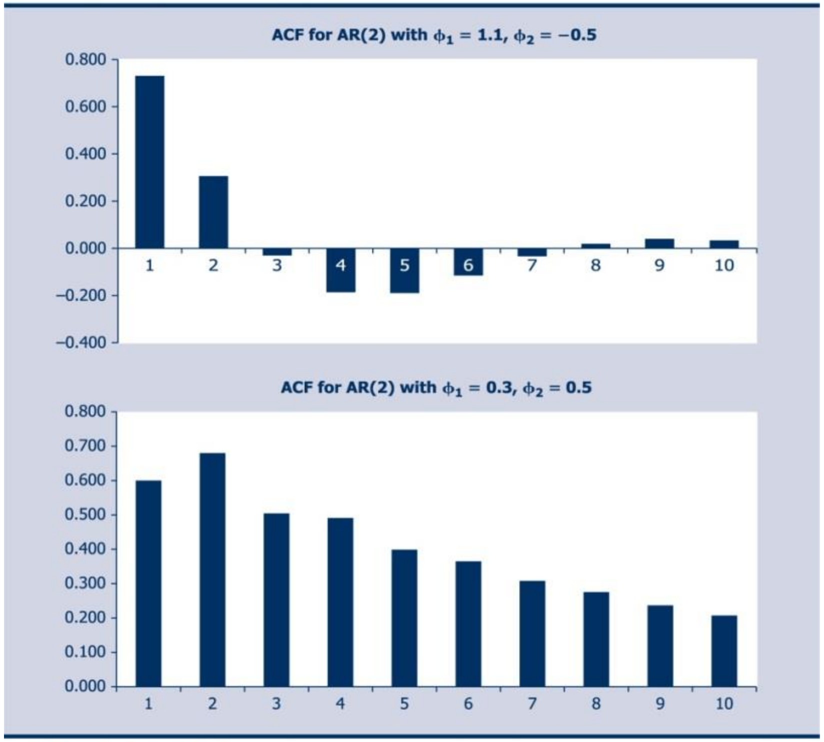
Q: Why do even-numbered lags have positive autocorrelation even when the coefficient is negative?

Sample ACFs will not look exactly like the theoretical forms!

Autoregressive Moving Average (ARMA) Models

- Higher order AR models have more complex ACF patterns.

FIGURE 6.3 ACFs FOR AN AR(2) SCHEME WITH (A) $\phi_1 = 1.1$ AND $\phi_2 = -0.5$ AND (B) $\phi_1 = 0.3$ AND $\phi_2 = 0.5$

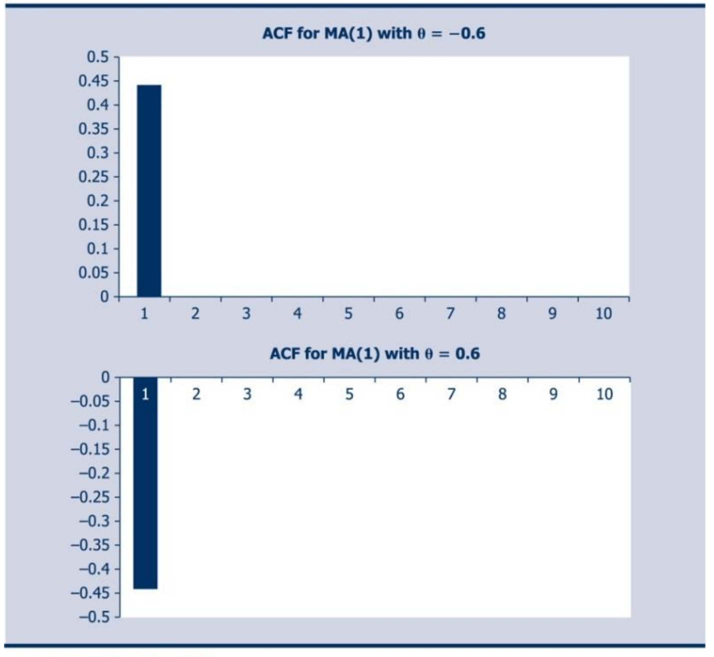


Q: What would the ACF look like when both coefficients ϕ_1 and ϕ_2 are negative?

Autoregressive Moving Average (ARMA) Models

- Moving Average models have a simple ACF structure.

FIGURE 6.4 ACF AND PACF FOR MA(1) SCHEME WITH (A) $\theta = -0.6$ AND (B) $\theta = 0.6$



Q: What would the ACF for an MA(2) scheme look like?

Partial Autocorrelations and Model Selection

- For k = 1,2, …, fit an AR scheme of order k to the data and record the value of the autoregressive coefficient for lag k. Denote this quantity by C_k.
- The plot C_k against k (=1, 2, 3, …) is the *Partial Autocorrelation Function* (PACF). For an AR(p) scheme, only the first p values will be nonzero.

FIGURE 6.6 PACF FOR AR(1) SCHEME WITH (A) $\phi = 0.8$, (B) $\phi = -0.8$ AND FOR MA(1) SCHEME WITH (A) $\theta = -0.6$, (B) $\theta = 0.6$

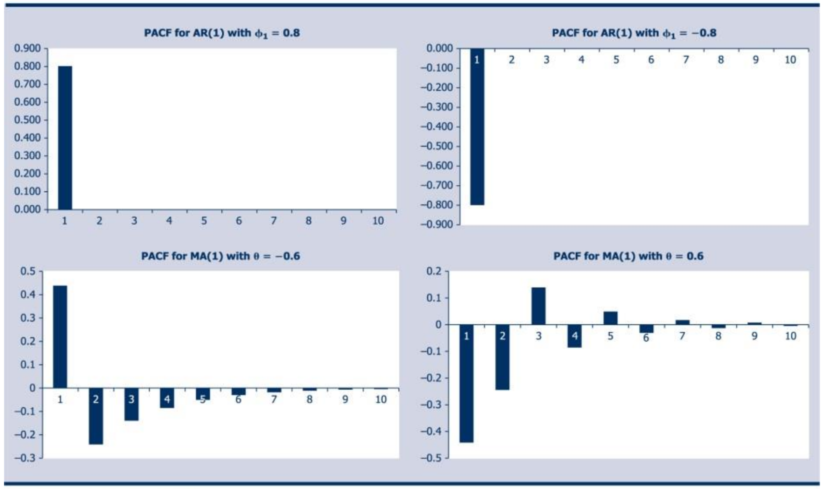
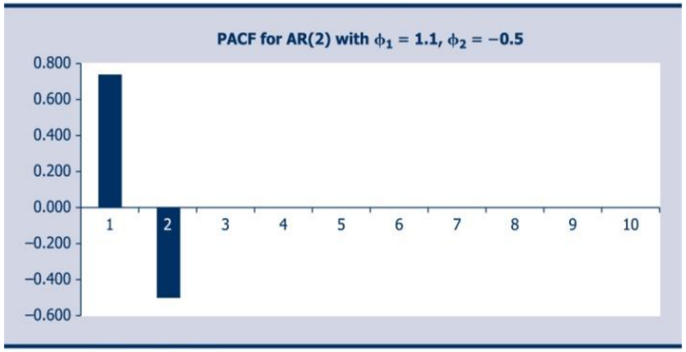


FIGURE 6.7 PACF FOR AR(2) SCHEME WITH $\phi_1 = 1.1$ AND $\phi_2 = -0.5$

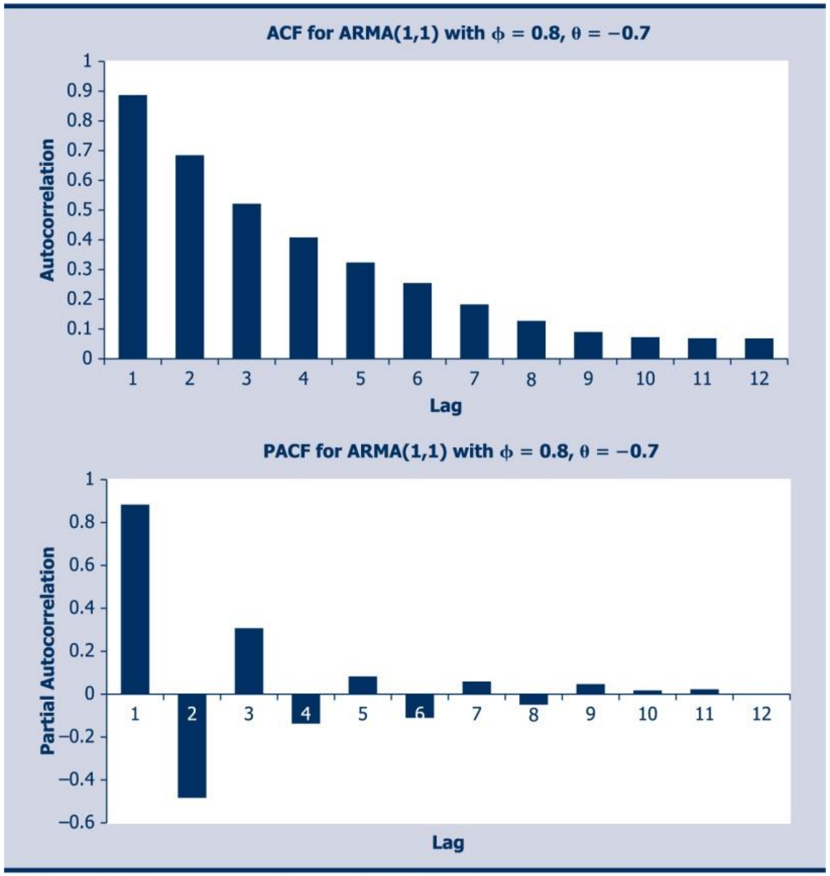


Partial Autocorrelations and Model Selection

- A model may have both autoregressive and moving average components. The ARMA(1,1) scheme has the form:

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

FIGURE 6.8 ACF AND PACF FOR AN ARMA(1, 1) SCHEME WITH $\phi = 0.8, \theta = -0.7$



Source: Based upon a generated series of 1000 observations.

Such mixed models have non-vanishing terms in both the ACF and the PACF.

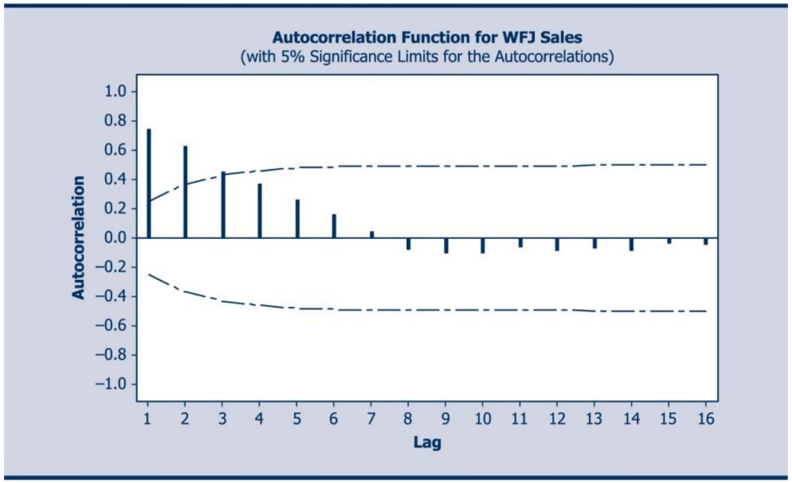
Partial Autocorrelations and Model Selection

TABLE 6.2 ACF AND PACF PATTERNS FOR GENERAL AR, MA, AND ARMA SCHEMES

Function	AR(p) Scheme	MA(q) Scheme	Mixed ARMA Scheme
ACF	Tails off as a damped wave pattern or damped exponential	Finite, q spikes	Tails off as a damped wave pattern or damped exponential
PACF	Finite, p spikes	Tails off as a damped wave pattern or damped exponential	Tails off as a damped wave pattern or damped exponential

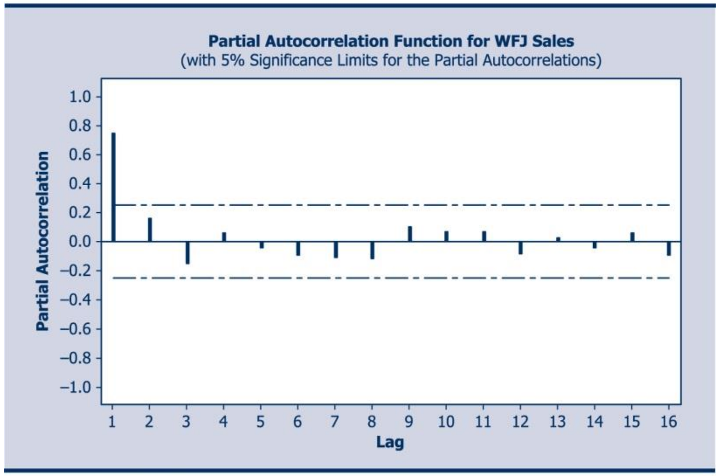
Partial Autocorrelations and Model Selection

FIGURE 6.1 THE ACF FOR WFJ SALES (MINITAB)



Data shown is from file WFJ_sales.xlsx.

FIGURE 6.5 THE PACF FOR WFJ SALES DATA



Data: WFJ_sales.xlsx.

Q: Which model would you suggest for WFJ Sales?

The golden rule is to be parsimonious:
pick the simplest model that seems to work.

Partial Autocorrelations and Model Selection

- When the series is non-stationary consider using a difference. The first order difference is:

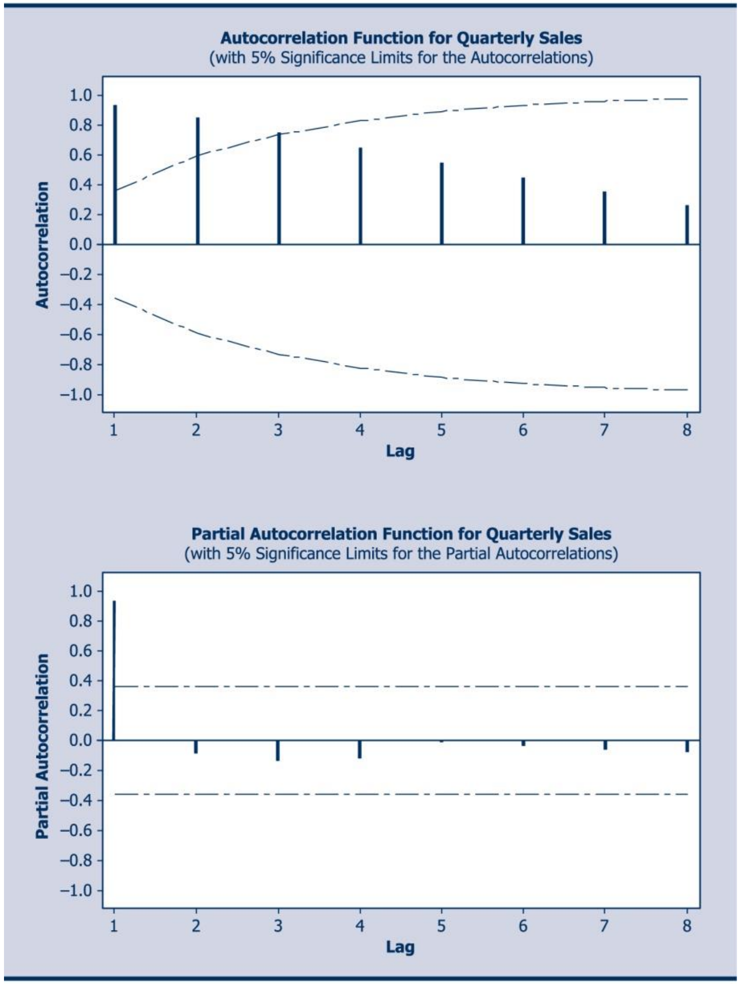
$$DY_t = Y_t - Y_{t-1}$$

- Converting to "Growth" is another option

- The ACF and PACF for Netflix sales are typical of non-stationary series [Netflix.xlsx]

- Note that the PACF displays the "Pollyanna effect"

FIGURE 6.10 ACF AND PACF FOR NETFLIX QUARTERLY SALES



Data: Netflix.xlsx.

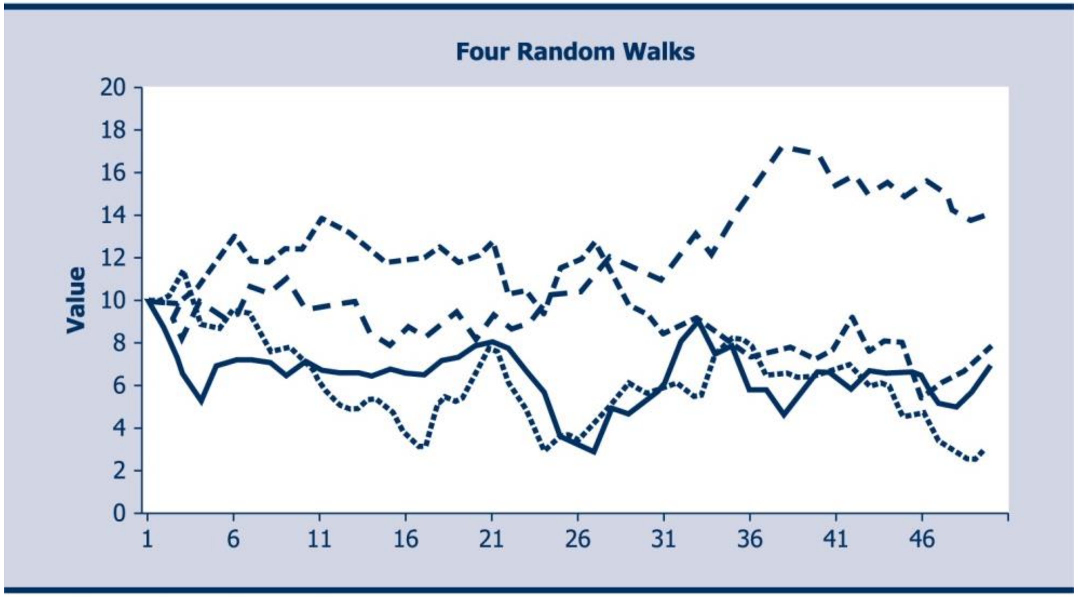
The Random Walk

- The Random Walk is defined by the simple model:

$$Y_t - Y_{t-1} = \varepsilon_t$$

- That is, the best forecast for the next period is the previous observation

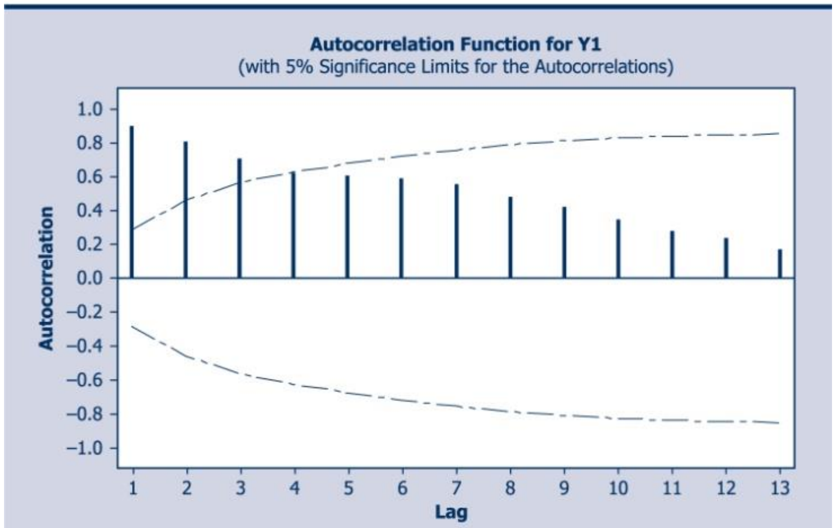
Four Random Walks



Q: The random walk is the standard model for stock prices. Why?

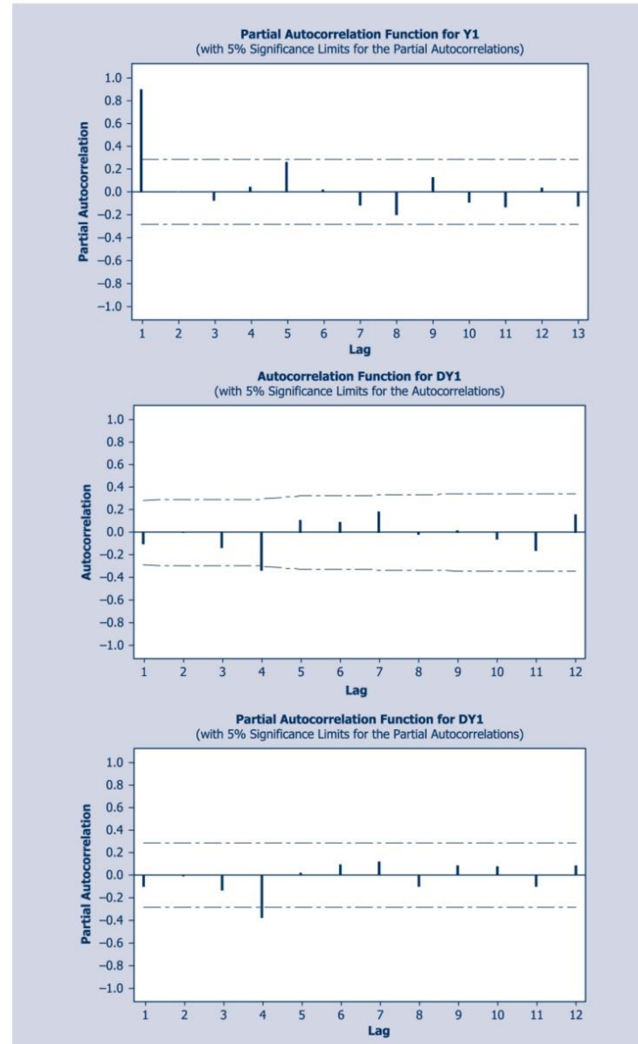
ACF and PACF for Random Walk and its First Difference

FIGURE 6.14 ACF AND PACF FOR A RANDOM WALK WITH $n = 50$, BEFORE AND AFTER DIFFERENCING



(Continued)

FIGURE 6.14 (CONTINUED)



Model Estimation and Selection

- Use the initial data plot, the ACF and the PACF to suggest possible models
- Examine the performance of the several models selected

TABLE 6.3 FITTED MODELS FOR NETFLIX PERCENTAGE GROWTH SERIES USING MINITAB

AR(1) Model

Type	Coef	SE Coef	T	P
AR1	0.6654	0.1423	4.68	0.000
Constant	5.101	1.665	3.06	0.005
Mean	15.243	4.976		

Mean Square error = 85.94 with DF = 29

[RMSE = 9.27]

MA(1) Model

Type	Coef	SE Coef	T	P
MA1	−0.4985	0.1603	−3.11	0.004
Constant	14.660	2.667	5.50	0.000
Mean	14.660	2.667		

Mean Square error = 98.77 with DF = 29

[RMSE = 9.94]

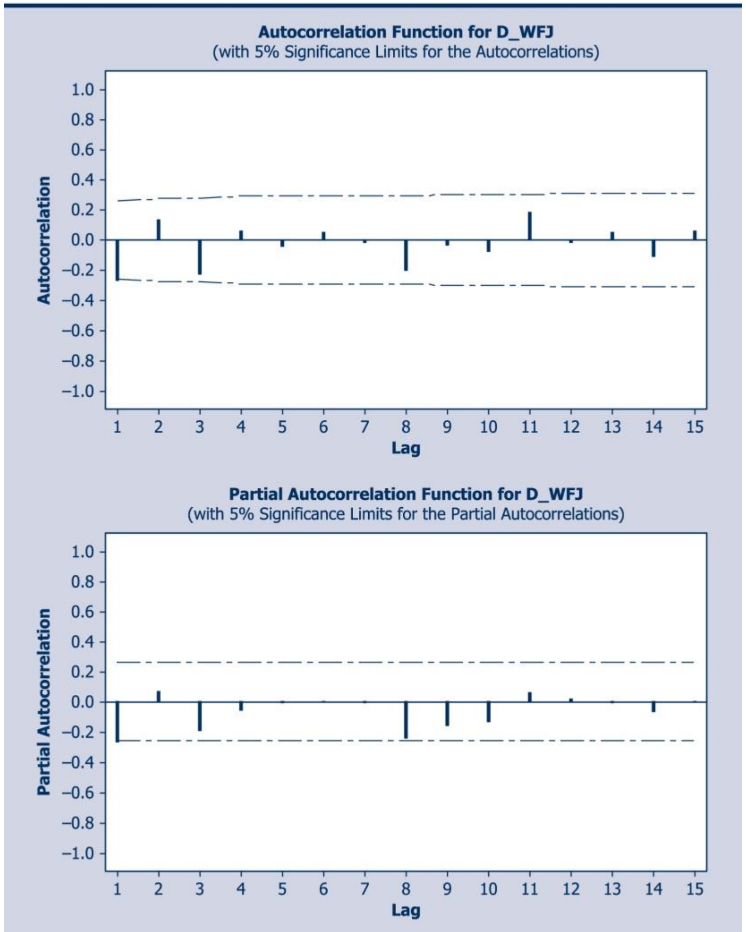
Data: Netflix.xlsx.

Q: Choose the better model. Which criterion would you use?

Model Estimation and Selection

Example 6.6: Fitted Model for WFJ Sales

FIGURE 6.15 ACF AND PACF FOR WFJ SALES DATA, DIFFERENCED ONCE



Data: WFJ_sales.xlsx.

TABLE 6.4 DIFFERENT MODELS FOR WFJ SALES DATA USING MINITAB

AR(1)

Type	Coef	SE Coef	T	P
AR1	0.7867	0.0807	9.75	0.000
Constant	6.8382	0.4460	15.33	0.000
Mean	32.063	2.091		

Mean Square error = 12.304 with DF = 60

[RMSE = 3.51]

ARIMA(1, 1, 0)

Type	Coef	SE Coef	T	P
AR1	−0.2766	0.1267	−2.18	0.033

Mean Square error = 12.933 with DF = 60

[RMSE = 3.60]

ARIMA(0, 1, 1)

Type	Coef	SE Coef	T	P
MA1	0.2686	0.1269	2.12	0.038

Mean Square error = 13.017 with DF = 60

[RMSE = 3.61]

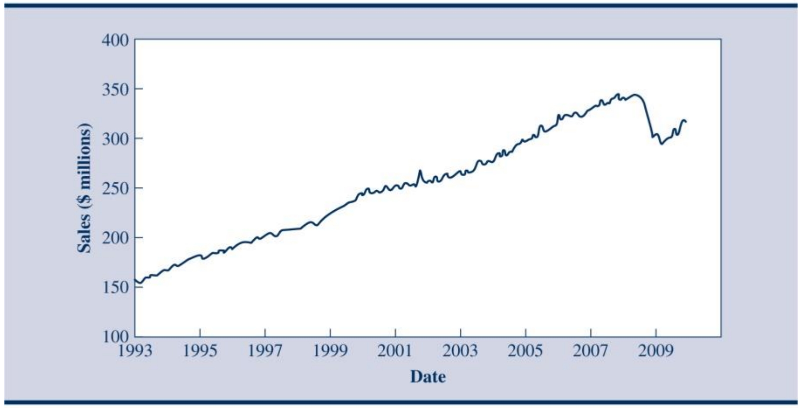
Data: WFJ_sales.xlsx.

Q: Which Model?

Model Estimation and Selection

Fitted Model for U.S. Retail Sales

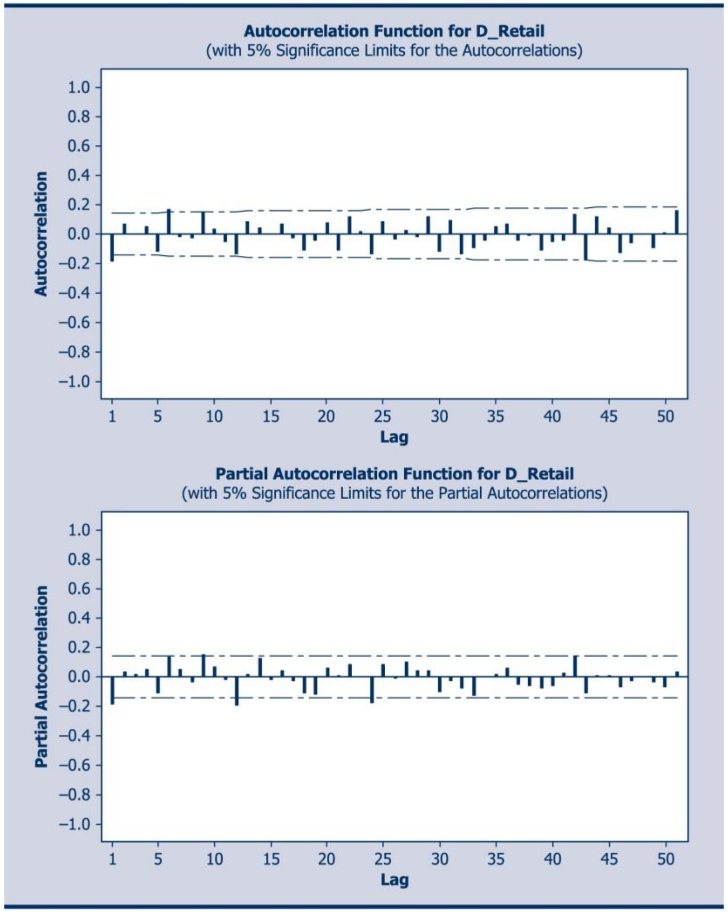
FIGURE 1.4 SEASONALLY ADJUSTED SERIES FOR U.S. MONTHLY TOTAL SALES FOR RETAIL AND FOOD SERVICES



Source: U.S. Census Bureau via EconStats, *www.econstats.com/retsal/retsal_m1.htm*.
Data: US_retail_sales.xlsx.

Q: Should we difference?
Which model would you suggest?

FIGURE 6.16 ACF AND PACF FOR FIRST DIFFERENCES OF U.S. RETAIL SALES (SEASONALLY ADJUSTED)



Data: US_retail_sales.xlsx.

Use of Information Criteria

- Use either $AIC = \ln(MSE) + 2p/n$

  or

  $BIC = \ln(MSE) + p\ln(n)/n$
- p = NP =number of parameters

Q: Which criterion is likely to select the more complex model? Why?

TABLE 6.7 VALUES OF INFORMATION CRITERIA FOR ARIMA MODELS FOR U.S. RETAIL SALES

		ARIMA				
Model	**Constant**	**Sample size**	**NP**	**RMSE**	**AIC**	**BIC**
1,1,0	no	202	2	2949	16.00	16.03
1,1,0	yes	202	3	2793	15.90	15.95
0,1,1	no	202	2	2952	16.00	16.03
0,1,1	yes	202	3	2797	15.90	15.95
2,1,0	yes	202	4	2799	15.91	15.98
0,1,2	yes	202	4	2798	15.91	15.98
1,1,1	yes	202	4	2799	15.91	15.98
0,2,2	no	201	3	2789	15.90	15.95

Data: US_retail_sales.xlsx.

Model Diagnostics

- What can go wrong?
  - The errors may not be normally distributed.
  - The data may contain outliers.
  - The errors may be autocorrelated.
  - The time series may be non-stationary.
  - The errors may show changing variances over time.
  - The mean of the errors may be non-zero (only applies when the model does not contain a constant term).
- How do we check for these problems and what do we do about them?

Model Diagnostics

TABLE 6.8 INTERPRETATION OF DIAGNOSTICS FOR RESIDUALS FROM ARIMA MODELS

Diagnostic Plots	Issues to Be Addressed	Indication of Problems	Remedial Action to Be Taken
Normal probability plot (NPP) and histogram (H)	**1.** The residuals may not be normally distributed. **2.** Are there any outliers?	**1.** The residuals do not follow a straight line (NPP), or the histogram is not bell shaped (H). **2.** Individual points are far removed from the overall plot (NPP) or other observations (H).	**1.** Consider a transformation. **2.** Adjust outliers.
Residuals versus fitted values	**1.** Is there evidence of nonlinearity? **2.** Are there any outliers? **3.** Does the series show increasing or decreasing variance with the size of the observations (heteroscedasticity)? **4.** (*Only when the model does not contain a constant term*) Is the model consistent with a zero mean for the errors?	**1.** The plot shows curvature. **2.** Individual points are far removed from the overall plot. **3.** The plot forms an increasing or decreasing funnel shape. **4.** The residuals plot shows evidence of a nonzero mean.	**1.** Consider a transformation. **2.** Adjust outliers. **3.** Consider a transformation. **4.** Add a constant term to the model.

(Continued)

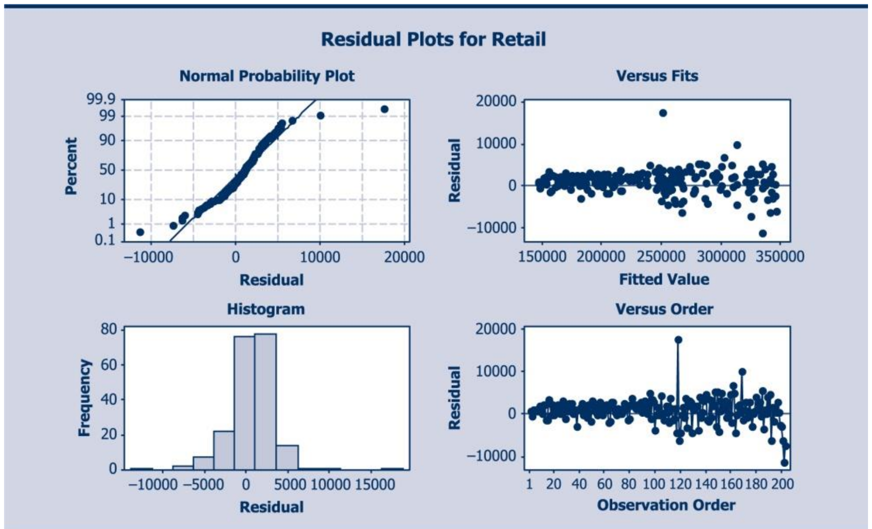
Model Diagnostics

TABLE 6.8 INTERPRETATION OF DIAGNOSTICS FOR RESIDUALS FROM ARIMA MODELS *(Continued)*

Residuals versus observation order	**1.** Is there any evidence of residual autocorrelation? **2.** Are there any outliers? **3.** Does the series show increasing or decreasing variance over time (changes in volatility)?	**1.** The plot shows long runs of residuals with the same sign (positive autocorrelation) or a zigzag pattern (negative autocorrelation). **2.** Individual points are far removed from the overall plot. **3.** The plot forms an increasing or decreasing funnel shape.	**1.** Consider adding terms to the model. **2.** Adjust outliers. **3.** Use a transformation to stabilize the variance.
Autocorrelation function (ACF)	**1.** Are additional differences required? **2.** Should more moving average terms be added? **3.** Is there too much differencing?	**1.** The ACF shows a slow linear decay. **2.** The ACF shows significant spikes. **3.** The lag 1 autocorrelation is close to its minimum value of −0.5.	**1.** Try further differencing. **2.** Include further MA terms. **3.** Reduce the amount of differencing.
Partial autocorrelation function (PACF)	**1.** Are additional differences required? **2.** Should additional AR terms be added?	**1.** A “Pollyanna” effect exists. **2.** The PACF shows significant spikes.	**1.** Try further differencing. **2.** Include further AR terms.

Model Diagnostics

Residual Plots for ARIMA (1,1,0) Model for U.S. Retail Sales

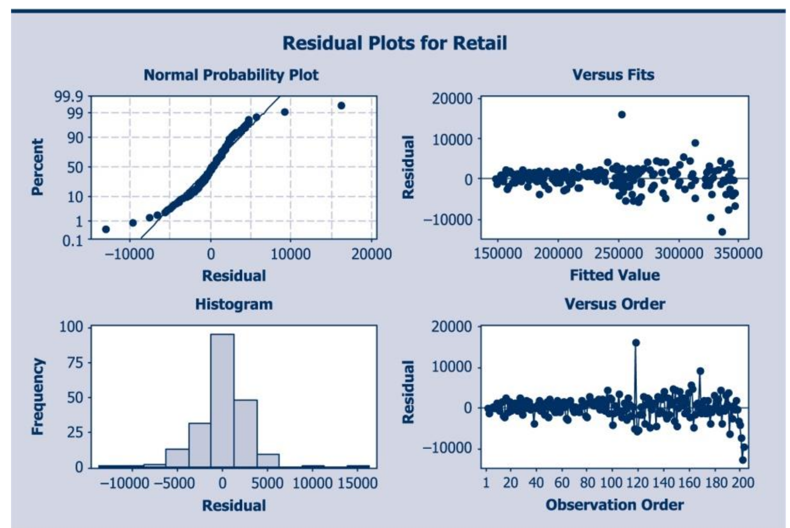


Data: US_retail_sales.xlsx.

ARIMA(0,1,1)

Figure 6.19: Residuals Plots for ARIMA (0,1,1)+C Model for U.S. Retail Sales

ARIMA(0,1,1) +C



Data: US_retail_sales.xlsx.

Q: What are the differences between the two sets of plots?

Model Diagnostics

The Ljung-Box statistic tests for the presence of autocorrelation:

$$Q = n(n+2)\sum_{k=1}^{k} \frac{r_k^2}{n-k}$$

An Application of the Ljung-Box Statistic to the U.S. Retail Sales Model

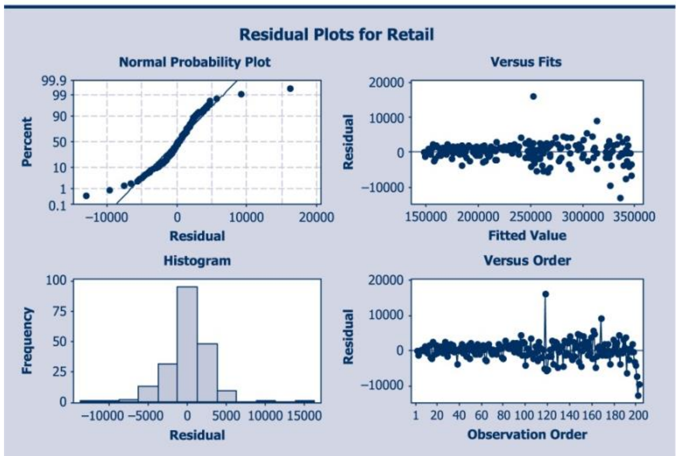
Modified Box-Pierce (Ljung–Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	18.9	35.2	53.2	78.9
DF	10	22	34	46
P-Value	0.042	0.037	0.019	0.002

Q: Is the model adequate for short-term forecasting?

Outliers Again

- After adjusting for the October 2001 outlier in the U.S. Retail Sales series, the diagnostics are as shown below. What conclusions may be drawn?

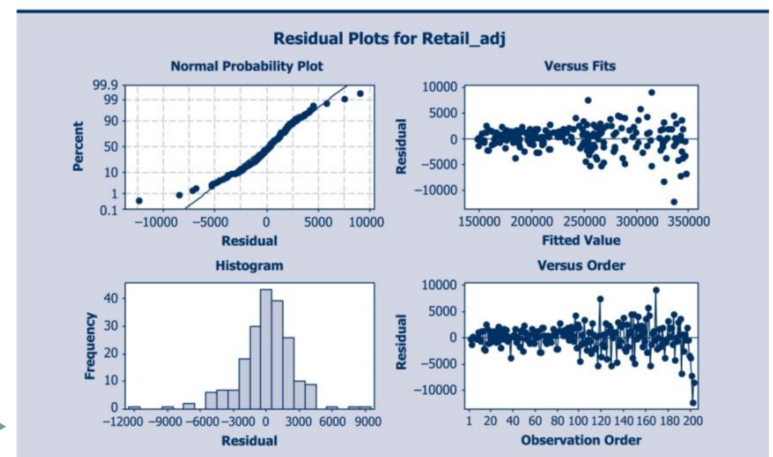
Residual Plots for ARIMA (0,1,1)+C Model for U.S. Retail Sales



Data: US_retail_sales.xlsx.

Original

Figure 6.20: Residuals Plots for U.S. Retail Sales, After Adjusting for the Outlier



Data: US_retail_sales.xlsx.

After removing outlier

Forecasting with ARIMA Models

Forecasting for WFJ Sales using AR(1) and ARIMA(0,1,1)

Y_t values: Use observations already recorded where available; for values not yet recorded, insert the latest available forecast.

ε_t values: Use recorded one-step-ahead residuals where available; set future errors in the equation equal to zero (best available forecast).

Model:

$$Y_{t+1} = 6.838 + 0.787 Y_t + \varepsilon_{t+1}$$

One-step-ahead forecast:

$$F_{63} = F_{63}(1) = 6.8382 + 0.7867 Y_{62}$$
$$= 6.8382 + 0.7867 * 34.1282$$
$$= 33.69.$$

Prediction Intervals

- In most cases, it suffices to compute the *h*-step-ahead mean square error and then construct the prediction interval as in earlier chapters.
- For stationary models, the long-term forecast approaches the mean and the MSE approaches a fixed value. For non-stationary series, the MSE increases without limit.

TABLE 6.9 FORECAST RMSE FOR WFJ SALES

Horizon, h	AR(1)	ARIMA(0,1,1)
1	3.51	3.61
2	4.47	3.74
5	5.42	4.10
10	5.66	4.64
20	5.69	5.56
50	5.69	7.69
σ	3.51	3.61
ϕ	0.7867	
θ		0.2686

Data: WFJ_sales.xlsx.

Forecasting Using Transformations

- Transform the series.
- Generate the point forecasts and prediction intervals for the transformed series.
- Convert the point forecasts and prediction intervals back to the original units.
- The prediction interval based upon a transformation will usually not be symmetric.

TABLE 6.10 POINT FORECASTS AND PREDICTION INTERVALS FOR NETFLIX SALES FOR 2008

Year	Quarter	LOG Model ARIMA(1,1,0)			ARIMA(1,1,0)+C		
		Forecast	Lower Limit 95% PI	Upper Limit 95% PI	Forecast	Lower Limit 95% PI	Upper Limit 95% PI
2008	1	322.4	276.7	375.7	311.3	295.2	327.3
	2	352.2	261.6	474.1	320.4	291.8	348.9
	3	390.9	252.1	606.3	329.6	290.2	369.1
	4	438.6	247.4	777.5	338.9	290.0	387.9

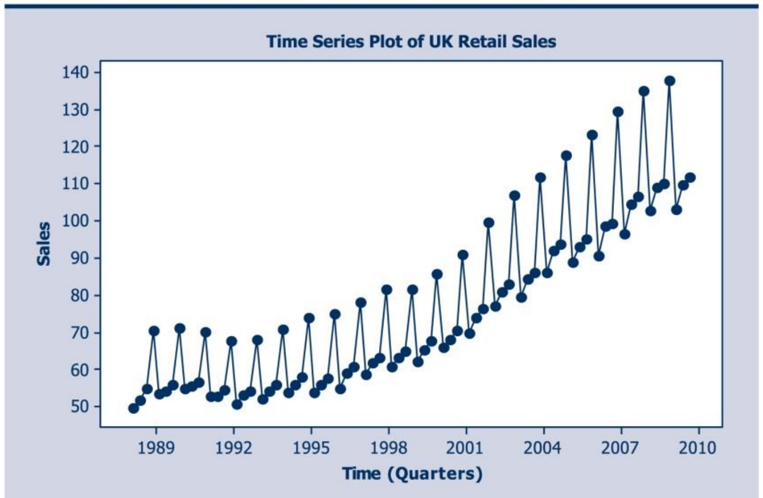
Data: Netflix.xlsx.

Seasonal ARIMA Models

- Combine regular and seasonal components.
- May need both regular and seasonal differences.
- The popular "airline model" is ARIMA$(0,1,1)(0,1,1)_m$.

Seasonal Model for UK Retail Sales

FIGURE 6.21 UK RETAIL SALES VOLUME, 1988Q1 TO 2009Q3



Data shown is from file UK_retail_sales.xlsx.

Type	Coef	SE Coef	T	P
MA 1	0.2186	0.1097	1.99	0.050
SMA 4	0.1183	0.1145	1.03	0.305

Differencing: 1 regular, 1 seasonal of order 4
Number of observations: Original series 87, after differencing 82
Residuals:

RMSE = 1.53 DF = 80

Modified Box-Pierce (Ljung–Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	15.6	24.2	30.6	37.4
DF	10	22	34	46
P-Value	0.113	0.338	0.635	0.813

Q: What are the potential weaknesses of this model?

State-Space and ARIMA Models

- We can express a *linear* state-space scheme in ARIMA form if and only if we assume that the random error terms satisfy the standard assumptions.
- Any ARIMA scheme may be expressed in state-space form.

TABLE 6.11 EQUIVALENCES BETWEEN STATE-SPACE AND ARIMA SCHEMES

State-space form	Parameter restrictions	ARIMA
Damped linear trend	None	(1,1,2)
Local linear trend (LES)	$\phi = 1$	(0,2,2)
Local level (SES)	$\phi = 0$	(0,1,1)

GARCH Models

- Generalized Autoregressive Conditional Heteroscedastic [GARCH] models provide descriptions of how the error variance evolves over time.

- The most commonly used model is the GARCH(1,1) scheme:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2.$$

- σ_t^2 denotes the variance at time *t* and ε_{t-1}^2 is the squared error at time *t-1*.

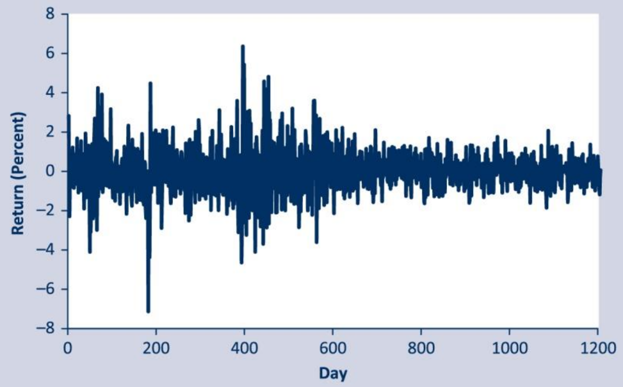
- $\omega,\ \alpha,$ and β are parameters subject to the conditions:

$$\omega > 0,\ \alpha > 0,\ \text{and } \beta > 0 \text{ and } \alpha + \beta < 1.$$

GARCH Models

Variation in the Dow Jones Index

FIGURE 6.24 DAILY RETURNS ON THE DOW JONES INDEX, JANUARY 2001–SEPTEMBER 2005 (*Analysis completed with EViews*)



Data shown is from file DowJones.xlsx.

TABLE 6.13 COEFFICIENT ESTIMATES FOR A GARCH(1,1) MODEL USING THE FIRST 1160 OBSERVATIONS OF THE DOW JONES SERIES (*Estimated with EViews 7*)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.028642	0.025845	1.108225	0.2678
		Variance Equation		
C	0.009478	0.004685	2.022987	0.0431
$RESID(-1)^2$	0.077483	0.011649	6.651471	0.0000
GARCH(−1)	0.915981	0.012865	71.19897	0.0000

Data: DowJones.xlsx.

Q: Interpret the coefficients.

Principles of ARIMA Modeling

- Identify the key features of the data.
- Transform the data to stationarity.
- Consider a variety of models, compatible with time series characteristics.
- Estimate the models, using your preferred software.
- Examine the diagnostic statistics-in particular, the residual autocorrelations.
- Consider simplifying the model if the coefficients are insignificant.
- Use an information criterion supplemented, where possible, by an out-of-sample fit to choose among models.
- Compare your ARIMA forecasts with alternative methods of forecasting - in particular, with a naïve model.

GARCH Models

- Start with a constant variance model, and then examine the autocorrelation structure of the squared residuals, using the methods of this chapter.
- The GARCH(1,1) model will suffice for most purposes.
- Remember KISS (Keep It Simple, Statistician)!

Take-Aways

- The Box-Jenkins or ARIMA approach provides a framework for model selection, estimation and forecasting.
- The ACF and PACF are the primary tools for model selection and they are also useful as diagnostic devices.
- ARIMA models require the time series to be stationary or to be transformed to stationarity. Typically business series will require differencing and/or transformations to achieve this condition.
- There is a general equivalence between *linear* state-space models and ARIMA models. However, the actual models selected in an application may well differ.
- GARCH models provide for changing variances to complement the changes in the mean level.

Reference and sourse:

1. Multivariate Time Series Analysis: With R and Financial Applications by Ruey S. Tsay
2. Time Series Analysis by James Douglas Hamilton
3. The Analysis of Time Series: An Introduction with R Chapman & Hall/CRC Texts in Statistical Science)
4. Machine Learning for Time Series Forecasting with Python by Francesca Lazzeri
5. Time Series Analysis for the Social Sciences (Analytical Methods for Social Research) Part of: Analytical Methods for Social Research (14 Books)
6. Introduction to Probability, Statistics, and Random Processes by Hossein Pishro-Nik
7. Introduction to Time Series and Forecasting (Springer Texts in Statistics) Part of: Springer Texts in Statistics