Time Series Regression - Part III

STAT 1321/2320

Kiran Nihlani

Department of Statistics
University of Pittsburgh

Outline

Differencing

- Order 1 Differencing (useful for non-seasonal data)

$$\nabla X_t = X_t - X_{t-1}$$

- Occasionally second-order differencing is required

$$\nabla^2 X_t = \nabla(\nabla X_t) = \nabla X_t - \nabla X_{t-1} = X_t - 2X_{t-1} + X_{t-2}$$

- We may also use seasonal differencing. $(x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$

The Backshift Operator

lag operator:

- Define the backshift operator B as

$$BX_t = X_{t-1}$$

$B(BX_t) = X_{t-2}$

- Similarly, $B^2 X_t = X_{t-2}$, and in general,

$$B^k X_t = X_{t-k}$$

- The inverse of the backshift operator is the forward-shift operator B^{-1}, such that

$$B^{-1} X_{t-1} = X_t, \text{ and } B^{-1} B X_t = X_t.$$

$B^{-1} X_{t-1} = X_t.$

Backshift Algebra

1) If X_t and Y_t are 2 time series.

$$B(X_t + Y_t) = BX_t + BY_t$$
$$= X_{t-1} + Y_{t-1}$$

2) If $X_t = c$, c : constant.

$$BX_t = X_{t-1} = c$$

3) c : constant

$$B(cX_t) = c\,BX_t = cX_{t-1}$$

① $X_t = \mu + 0.6X_{t-1} + w_t$

$= \mu + 0.6BX_t + w_t$

② $X_t = w_t + 0.6w_{t-1} + 0.3w_{t-2}$: $MA(2)$

$= [1 + 0.6B + 0.3B^2]w_t$

③ $X_t = 0.6X_{t-1} + w_t + 0.4w_{t-1}$

$X_t - 0.6X_{t-1} = w_t + 0.4w_{t-1}$

$(1 - 0.6B)X_t = (1 + 0.4B)w_t$

The Backshift Operator and Differencing

- The first difference operator ∇ can be expressed as

$$\nabla X_t = (1 - B)X_t$$

$x_t - x_{t-1}$

- The second difference is simply

$$\nabla^2 X_t = (1 - B)^2 X_t = X_t - 2X_{t-1} + X_{t-2}$$

- In general,

$$\nabla^d X_t = (1 - B)^d X_t$$

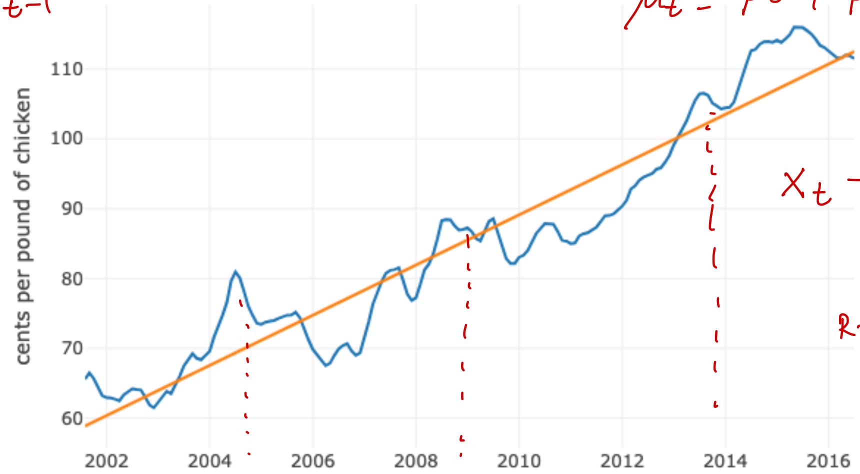
$$(1 - B)^2 x_t = (1 - 2B + B^2) x_t = x_t - 2x_{t-1} + x_{t-2}$$

Example - Chicken Prices

spot price. $X_t = \mu_t + \varepsilon_t$

$\mu_t = \beta_0 + \beta_1 t$.

$x_t - x_{t-1}$

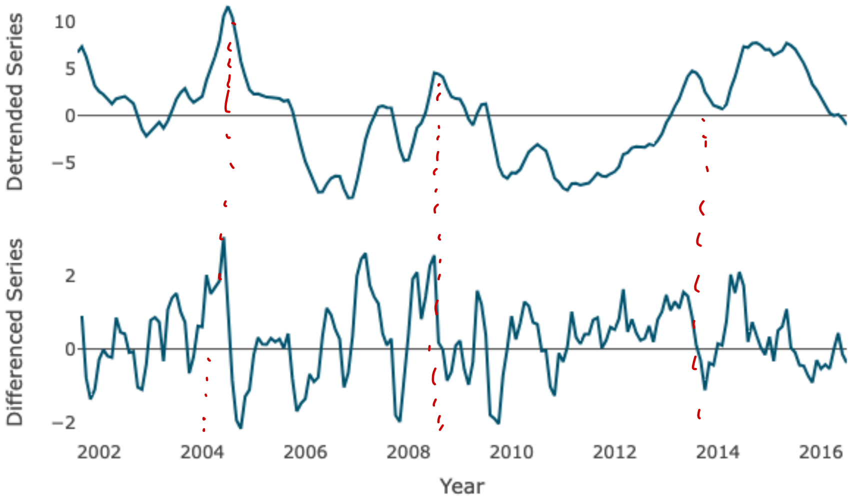


$X_t - \mu_t$

$= \varepsilon_t$

Residuals.

Chicken Price - Detrending vs Differencing



Chicken Price - Observations

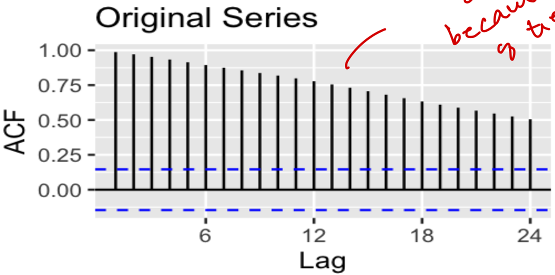
diff (x, order=)

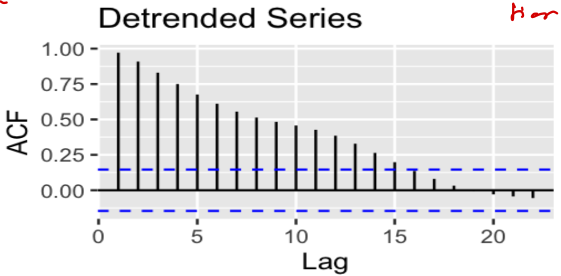
- The original series showed (sort of) linear, positive trend.
- The de-trended series (residuals from lm model) and differenced series (first order) are plotted:
  - The de-trended series is smoother and shows a 5 year cycle which the differenced series does not capture.
- The sample ACFs are plotted on the next slide:
  - The original and de-trended series show a slow decay that is reflective of the trend.
  - The differenced series shows probable 1 year cycle.

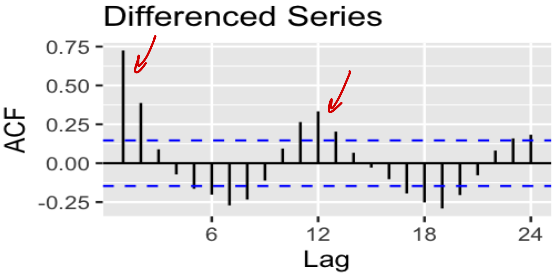
Chicken Price - ACF

slow decay because of trend.

residual. significant auto-correlation.

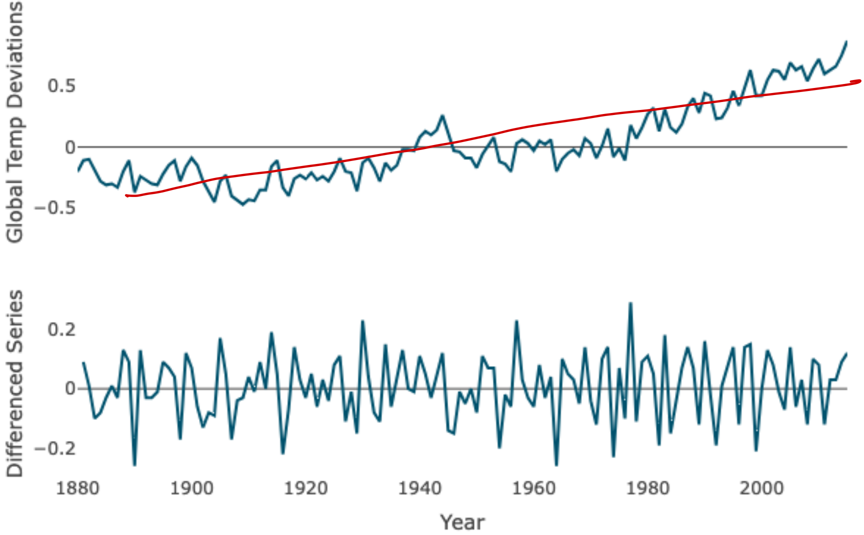




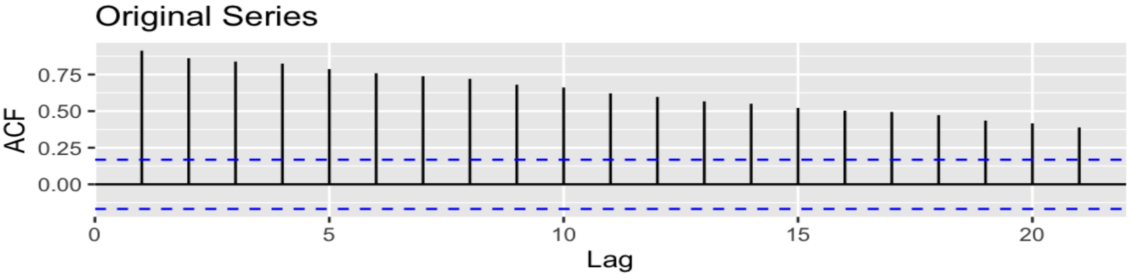


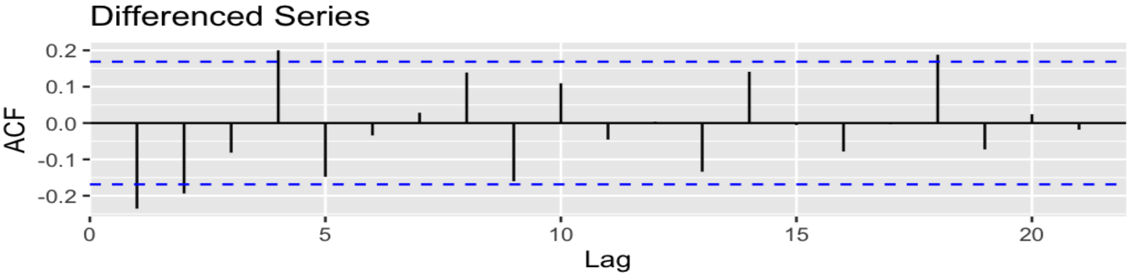
periodicity.

Example - Temperature Deviations



Example - Temperature Deviations





RW without drift:

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

$$= \sum_{i=1}^{t} w_i$$

with drift:

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

$$= \delta t + \sum_{i=1}^{t} w_i$$

$$\nabla X_t = X_t - X_{t-1}$$

$$= \delta t + \sum_{i=1}^{t} w_i$$

$$- \delta(t-1) - \sum_{i=1}^{t-1} w_i$$

$$= \delta + \sum_{i=1}^{t} w_i - \sum_{i=1}^{t-1} w_i$$

$$E[\nabla X_t] = \delta$$

Example - Temperature Deviations

- The global temperature deviations show an upward trend. Possible models?
- Consider the differenced series:
  - The differenced series plot looks like a stationary series (potentially a white noise).
  - The ACF for the differenced series shows minimal auto-correlation.
- This indicates that the temperature deviations data may be modeled with a radnom walk with drift

$$X_t = \delta + X_{t-1} + W_t = \delta t + \sum_{j=1}^{t} W_j$$

- The value of δ can be found as the mean of the differenced series ($\delta = 0.008$).

Random Walk

with drift.

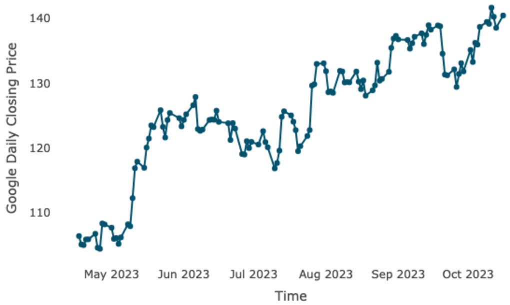
- Random walk models are widely used for non-stationary data, particularly financial and economic data.
- Random walks typically have:

  jagged.

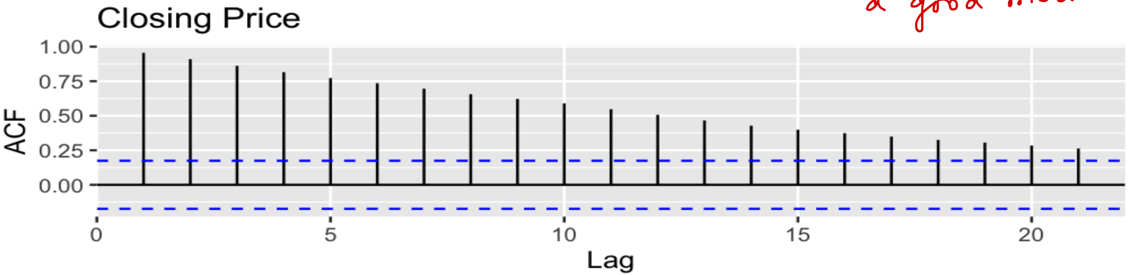
  - long periods of apparent trends up or down
  - sudden and unpredictable changes in direction.
- Example: for a stock price data random-walk theory asserts that there is no pattern to stock-price changes. ← w_t.
  - In particular, past stock-price changes do not enable one to predict future price changes.

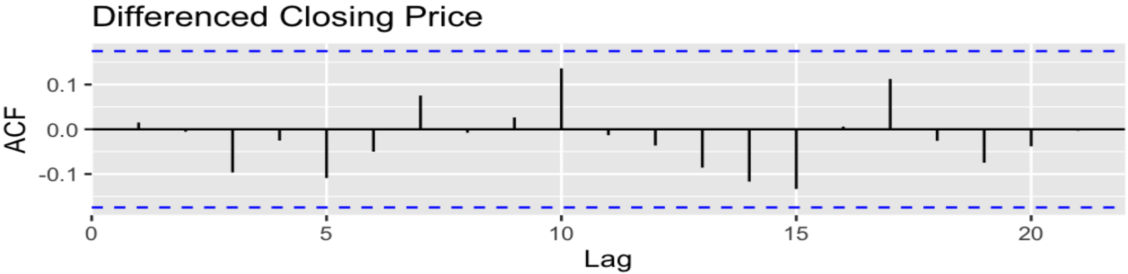
Example - Google

- The jagged appearance of the graph suggests that this may be a random walk.
- As the price change at one moment is uncorrelated with past price changes, the incessant up-and-down movement makes the graph jagged.
- What if the graph was smooth?



Example - Google → random walk with drift is a good model.





Notes

- Differencing can help stabilise the mean of a time series by removing changes in the level of a time series, and therefore eliminating (or reducing) trend and seasonality.
- Transformations such as logarithms can help to stabilise the variance of a time series.
- As well as looking at the time plot of the data, the ACF plot is also useful for identifying non-stationary time series.
  - For a stationary time series, the ACF will drop to zero relatively quickly, while the ACF of non-stationary data decreases slowly.
  - Also, for non-stationary data, the value of autoccorelation at lag 1 is often large and positive.

Seasonal Differencing

- A seasonally differenced series is obtained by taking the difference between an observation and the previous observation from the same season

  t

$$X_t - X_{t-m}$$

$(1 - B^m) x_t .$

  where m is the seasonality period.

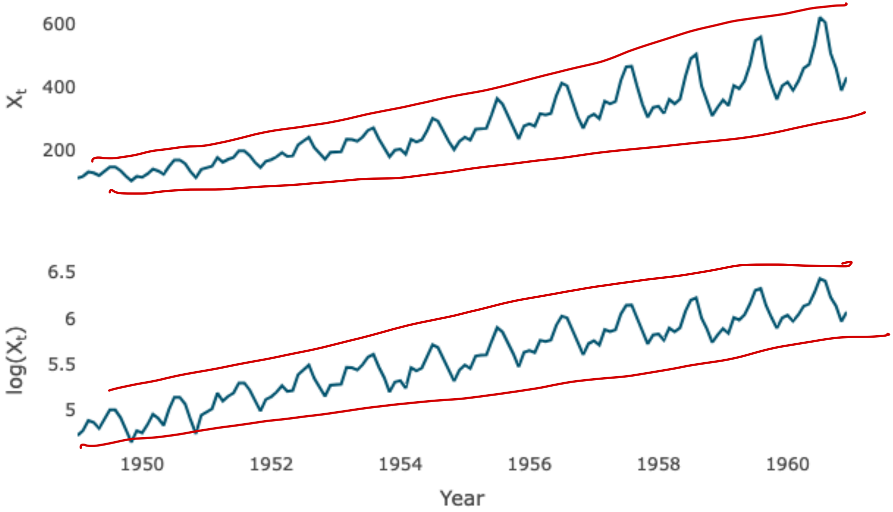
- These are also called "lag-m differences".
- If seasonally differenced data appear to be white noise, then an appropriate model for the original data is

$$X_t = X_{t-m} + \epsilon_t$$

$x_t - x_{t-m}$

$= \varepsilon_t .$

Example - Air Passenegers Data

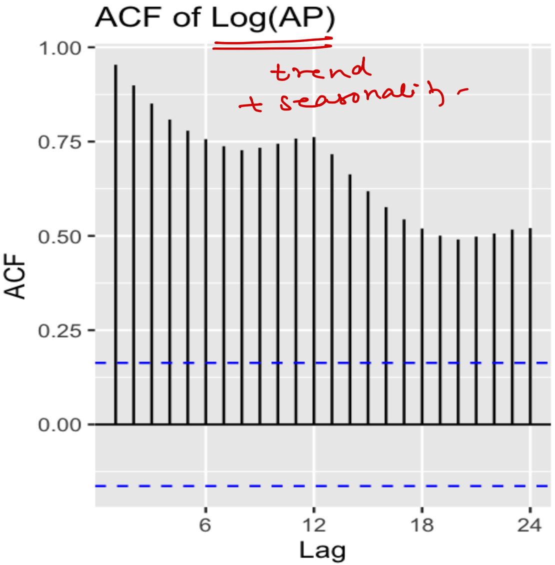


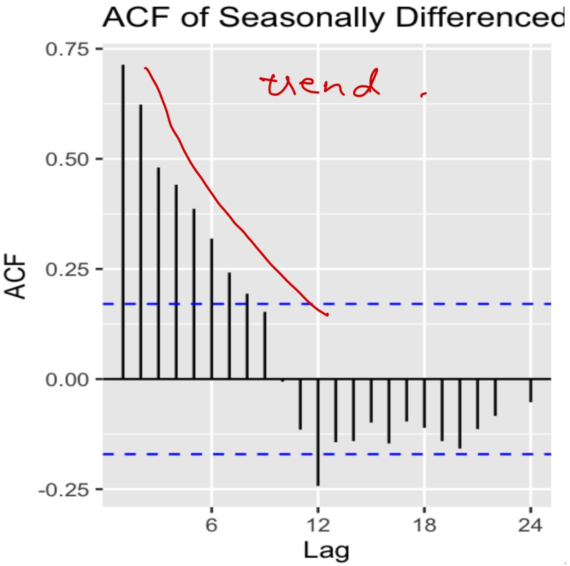
Example - Air Passenegers Data

- The log transform stabilized the variance.
- Seasonally differencing the series should help in accounting for the seasonality but is it enough to get a stationary series?
- Sometimes, the data may require a first order differencing along with seasonal differencing (if there is a dominant trend).
- In such cases, the order of differencing is not important but it is preferred to start with seasonal differencing if there is a strong seasonal pattern.
- The big question - when should we stop differencing?
- Plot the ACFs at every step to see if the resulting series looks stationary.

Example - Air Passenegers Data

$x_t - x_{t-12}$

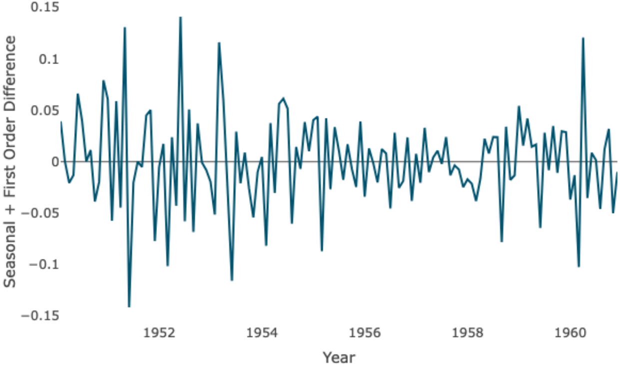


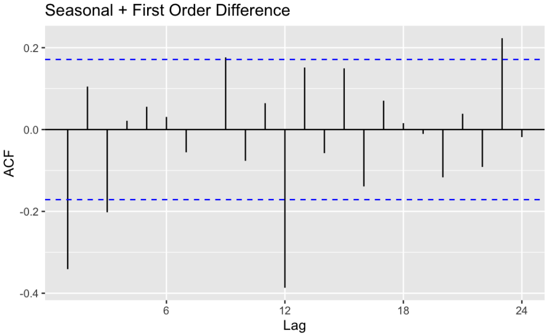


Example - Air Passenegers Data

$x_t - x_{t-12} = y_t$

∇y_t





Differencing

$x_t - x_{t-1}$

- It is important that if differencing is used, the differences are interpretable.
- First order differences represent the change between subsequent observations.
- Seasonal differences represent changes from one year to the next.
- Other lags are unlikely to make much interpretable sense and should be avoided.