Detecting a Non-Stationary Mean

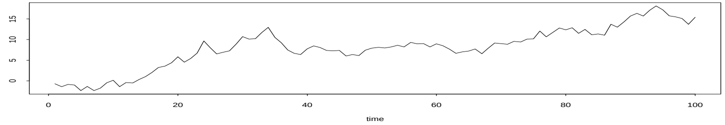
Visual inspection of the time series plot will reveal many types of non-stationarity.

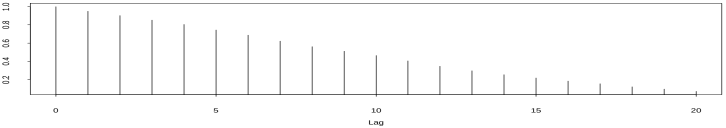
For detecting non-stationary mean, the Sample ACF (SACF) is also helpful.

- A realization from a process with a NON-stationary mean will typically have a SACF which decays very slowly to zero. (Note: the theoretical ACF of a process with non-stationary mean does NOT exist.)
- This fact may be used to help recognize such processes.
- Examples: Realizations from non-stationary AR(1) and AR(2) processes.

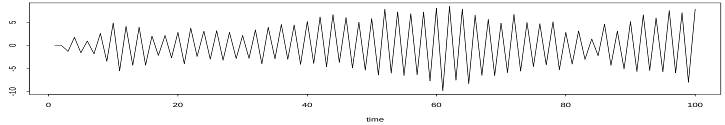
Sample ACF's for Realizations of Non-stationary Processes

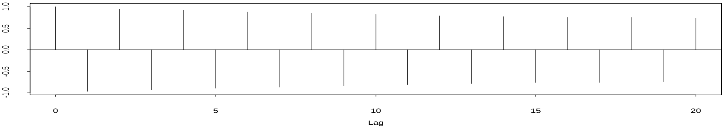
Random Walk: AR(1) with phi(1) = 1.0





Nonstationary AR(2) with phi(1) = -.5 and phi(2) = +.5





Warnings:

- Other types of non-stationarity (such as non-stationary variance) do not show up on the SACF. So, inspection of the time series plot is still required.
- If the autocorrelations of a series are changing with time, that also will **not** show up on the SACF. But if you divide the series into parts, and compare the SACF for each part, then it may show up.
- It is possible for a stationary ARMA process to have an ACF which decays slowly. Thus it will sometimes be hard to distinguish between nonstationary and stationary processes. Example:
  - The theoretical ACF of the AR(1) process is $\rho_k = \phi_1^k$
  - With $\phi_1 = .9$ we have $\rho_1 = .9, \rho_2 = .81, \rho_3 = .729, \ldots$ which is decaying rather slowly to zero.
  - Also, the SACF for a realization of this processs will typically decay to zero slowly.

- Thus, it is difficult in practice to distinguish between the non-stationary AR(1) with $\phi_1 = 1.0$ and the stationary AR(1) with $\phi_1 = 0.9$.
- Many other similar examples can be constructed.

- Given a particular realization, deciding whether the underlying process is stationary or not can sometimes be very difficult. In close cases, the decision is highly subjective. The decision is often made on the basis of simplicity. (Which choice leads to a simpler model?)

ARIMA processes

• ARIMA processes are generated (simulated) by "integrating" ARMA processes. The "I" in ARIMA stands for "integrated".

Suppose we have a series $z_1, z_2, \ldots, z_n$.

• "Integrating" a time series means computing *cumulative sums* for that series.

When we say that "$\{y_t\}$ is obtained by integrating $\{z_t\}$" we mean that

- $y_t = z_1 + z_2 + \cdots + z_t$ for all times $t = 1, \ldots, n$.
- A mathematically equivalent way to state this is

$$y_t = y_{t-1} + z_t \quad \text{for } 2 \leq t \leq n$$

where we define $y_1 = z_1$.

• "Differencing" a series means computing the differences between consecutive values of the series.

When we say that "$\{x_t\}$ is obtained by differencing $\{z_t\}$" we mean that

$$x_t = z_t - z_{t-1} \quad \text{for } t = 2, 3, \ldots, n\,.$$

We cannot compute x_1; this value is missing (indicated by a period "." in SAS).

If we difference a series d times, we end up with a series with d missing values at the beginning.

The symbol ∇ (nabla) is often used to denote differencing:

$$\nabla z_t = z_t - z_{t-1}\,.$$

Using backshift notation, we may write $\nabla = 1 - B$ since

$$\nabla z_t = z_t - z_{t-1} = z_t - Bz_t = (1 - B)z_t\,.$$

• "Differencing" is almost the opposite (inverse operation) of "integrating". If you integrate and then difference, you get back the original series except for a missing value at the beginning.

• An ARIMA(p,d,q) process is obtained by integrating an ARMA(p,q) process d times. (The quantity d is referred to as the "degree of differencing", p is the AR order, q is the MA order.)

• Differencing an ARIMA(p,d,q) process d times produces an ARMA(p,q) $\equiv$ ARIMA(p,0,q) process.

• Integrating a stationary ARMA process produces a **non**-stationary ARIMA process.

• Here are some time series plots of ARIMA processes.

Remarks on Differencing

- Purpose of Differencing: If a process (or realization) has a nonstationary mean, differencing one or more times may produce a stationary process (or realization).
- In particular, if the original process is ARIMA(p, d, q), then differencing d times produces an ARMA(p, q) process (which will be stationary if the AR coefficients satisfy the stationarity conditions).
- For series typically encountered in practice, a degree of differencing of $d = 0$, 1, or 2 is generally adequate. The values $d = 0$ (no differencing) and $d = 1$ are quite common. The value $d = 2$ occurs less frequently.
- If your series seems to require $d > 2$, probably another approach is needed.
- Differencing may eliminate a non-stationary mean, but it will *not* cure other forms of non-stationarity, such as nonstationary variance or nonstationary ACF.

Stabilizing the Variance

- Many time series exhibit the following behavior: The variability of the series z_t changes systematically with the level of the series.
  - Many economic time series have a long term positive trend due to economic growth or inflationary effects. For such series it is common for the variability to increase as the series climbs. Frequently, the standard deviation is roughly proportional to the value of the series; when the series doubles, the standard deviation doubles.
- This is a form of nonstationary variance.
- This problem can frequently be "cured" by transforming the data. We model the series $y_t = f(z_t)$ for some appropriately chosen function f, instead of modeling the original series z_t.
  - If the variability (as measured by the "local standard deviation") is *proportional* to the level of the series z_t, the appropriate transformation is taking logs, that is, we model $y_t = \log(z_t)$.

- If the variability (standard deviation) increases at a slower rate, a square root transformation ($y_t = \sqrt{z_t}$) may work.
- More general power transformations $y_t = z_t^{\lambda}$ also may be employed. (Here the "power" is λ.) The value of λ is chosen (perhaps by trial and error) to stabilize the variance.
  - A convenient form of the power transformation is the so-called Box-Cox transformation

    $$y_t = \frac{z_t^{\lambda} - 1}{\lambda}.$$

    For practical purposes, this is essentially the same as the simple power transformation $y_t = z_t^{\lambda}$.
- Sometimes it is helpful to add or subtract a constant before using a log or power transformation, that is, we model $y_t = \log(z_t - c)$ or $y_t = \sqrt{z_t - c}$, etc.

• If you obtain forecasts $\hat{y}_t$ for a transformed series $y_t = f(z_t)$, then you must apply the inverse transform $\hat{z}_t = f^{-1}(\hat{y}_t)$ to obtain forecasts for the original series.

Modeling a Time Series with Non-Stationary Mean

- Box-Jenkins approach: Model the series as a realization of an ARIMA(p, d, q) process. Choose d to be the order of differencing needed to make the series stationary. Then choose p and q by finding an ARMA(p, q) model for the differenced series. (Do this by studying the ACF and PACF of the differenced series.)
- Alternative: Model the series as

$$(\text{Series}) = (\text{Trend}) + (\text{Stationary Process})$$

  The stationary process could be an ARMA process. The trend could be linear $(a + bt)$, quadratic $(a + bt + ct^2)$, sinusoidal $(\sin(\omega(t - t_0))$, a periodic function represented as a sum of sines and cosines, etc. Here t denotes time.

Identifying ARIMA Models

Choose the order of differencing d:

- If the original series $z_1, \ldots, z_n$ appears stationary (e.g., the time series plot has a constant mean and the SACF decays fairly rapidly), then try $d = 0$.
- If not, examine the first differences $y_t = (1 - B)z_t$. If they appear stationary, try $d = 1$. ($d = 1$ is very common.)
- If not, examine the second differences $w_t = (1 - B)^2 z_t$. If they appear stationary, try $d = 2$. (This is less common.)
- If not, try something else. ($d > 2$ is rare.)

Overdifferencing

How can I tell if I have differenced too many times?

The **Inverse Autocorrelation Function (IACF)** may be useful here.

What is the IACF?

The theoretical IACF of an ARMA(p,q) process is equal to the theoretical ACF of the corresponding ARMA(q,p) process obtained by interchanging the roles of the θ's and ϕ's.

In particular,

- The theoretical IACF of an AR(p) = ARMA(p,0) process is the same as the ACF of an MA(p) = ARMA(0,p) process; it has a cutoff (to zero) after lag p.
- The theoretical IACF of an MA(q) process is the same as the ACF of an AR(q) process; it eventually decays to zero (perhaps in a complicated way).

The IACF and PACF are different, however, their general interpretation is similar for nonseasonal stationary processes. For an AR(p) process, both the theoretical IACF and PACF cutoff to zero after lag p. For seasonal processes (discussed later), the IACF may be easier to interpret than the PACF.

The **sample** IACF of an "over-differenced" series will typically decay to zero very slowly. That is, if you take a realization from an already stationary ARMA process and difference it, the sample IACF of the resulting differences will usually decay to zero very slowly.

Thus the ACF and IACF are both of assistance in choosing the appropriate order of differencing, the ACF helping to detect "under-differencing" and the IACF helping to detect "over-differencing".

After choosing d ...

Now choose the AR and MA orders p and q in the usual way, by examining the sample ACF/PACF of the appropriately differenced series.

There may be a number of plausible models, perhaps even with different values of d (e.g., $d = 0$ or $d = 1$ may both be plausible for "borderline" stationary series). Choose among them by looking at the estimation results:

- Significance of estimated parameters
- Residual diagnostics: ACF, PACF, Autocorrelation Check of Residuals (the Chi-Square statistics, i.e., Ljung-Box Q), QQ-Plot, plot of residuals versus predicted values.
- AIC or SBC (Note: these should only be used to compare models having the same number of residuals.)

If the variability of $z_1, \ldots, z_n$ changes systematically with the level of the series, try a transformation (e.g., log, square root, other power transform, etc.). Transformations are always applied to the original series, never to the differenced series.

Sometimes one does not realize a transformation is needed until after fitting an ARIMA model to the raw data and observing non-constant variance in the plot of residuals versus predicted values.