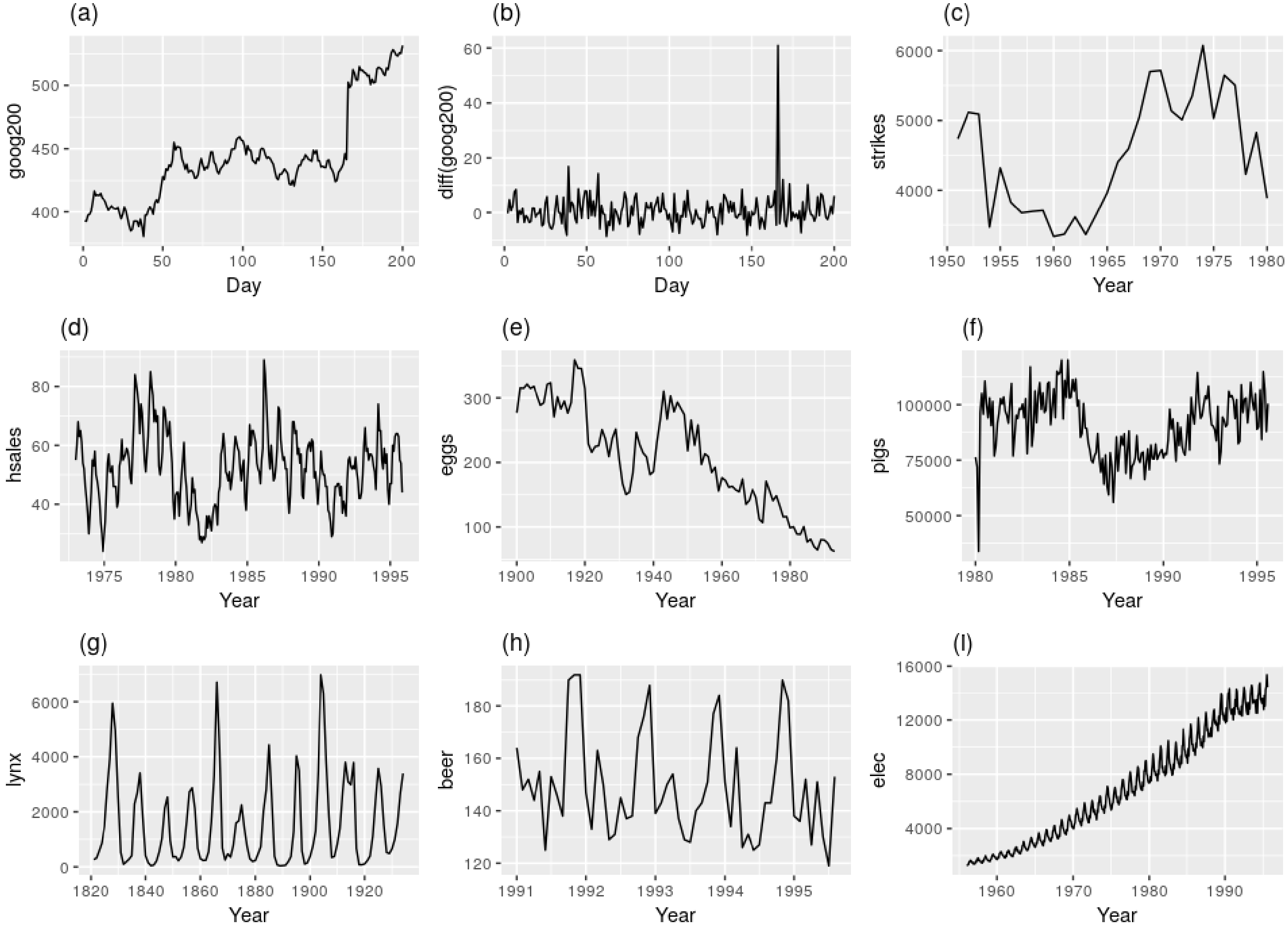
TSA04: Univariate Non-stationary Time Series Models

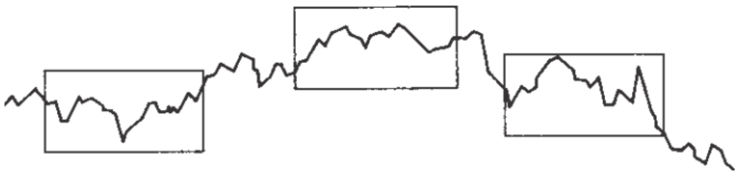
Jakey BLUE

https://otexts.com/fpp2/stationarity.html

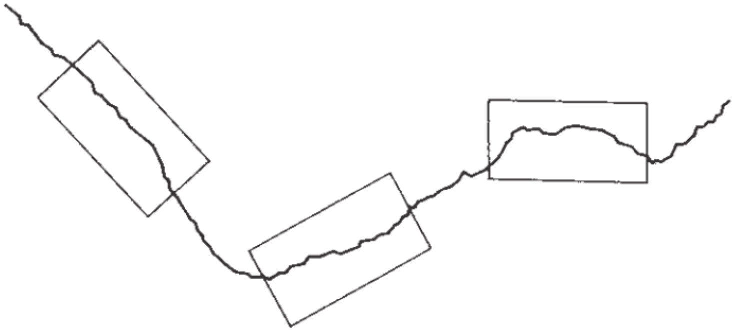


Time series is mostly non-stationary!

- AR, MA, ARMA are rigorous models for stationary processes. However, in real life, most processes are non-stationary, like these.
  - Homogeneous Nonstationary Time Series



$\nabla = 1 - B$



$\nabla^2 = (1 - B)^2$

Transform the Time Series into a Stationary One

A Generalized Form for Time Series Models

1

ARIMA: AutoRegressive Integrated Moving Average

∇: del or nabla

- Introducing the "differencing operator": $\nabla = 1 - B$
  - $z_t = \nabla y_t = y_t - y_{t-1}$
  - $y_t = S z_t, S = \nabla^{-1} = ?$

  S is usually called "infinite summation operator".

  - Think about: given z_t, how do you come back to the series y_t, if $z_t = \nabla y_t$.
    - Moreover, if $z_t = \nabla^2 y_t$, what's y_t in terms of z_t

Properties/Limitations of ∇^d

- ∇ is not the only operator for stationary transformation
  - $\sqrt{y_t}$ or $\ln(y_t)$ will work.
  - ∇ is in the sense of "slope". ∇^2 is similar to the "acceleration".
  - ∇ shorten the length of series, ∇^d will shrink the length to be $n - d$
  - When d of ∇^d is too large (e.g., ≥ 3), the series will easily "get lost" and become meaningless.

How do we know d?

ARIMA (p, d, q) Process

- $\phi(B)\nabla^d y_t = \theta(B)a_t + L,$
  - where
    $a_t \sim^{iid} N(0, \sigma_a^2)$;
    $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$;
    $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$;
    L is constant (and usually ignored as the original series is centered).
  - The roots of $\phi(B)$ and $\theta(B)$ are all outside the unit circle ➔ stationarity and invertibility.
  - ARIMA generalizes ARMA, ARI, IMA.

Three Explicit Forms of ARIMA

- Difference Equation Form

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_{p+d} y_{t-p-d} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

- Random Shock Form

$$y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \cdots$$

- Inverted Form

$$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \pi_3 y_{t-3} + \cdots + a_t$$

ARIMA(1, 1, 1) in Difference Equation Form

- $(1-\phi B)(1-B)y_t = (1-\theta B)a_t$

  $[1-(\phi+1)B+\phi B^2]y_t = a_t - \theta a_{t-1}$
  $y_t = (\phi+1)y_{t-1} - \phi y_{t-2} + a_t - \theta a_{t-1}$

  - Something familiar? Recall ARMA(2, 1) $\Rightarrow y_t = \phi_1 y_{t-1} - \phi_2 y_{t-2} + a_t - \theta a_{t-1}$

    - $\phi_1 = (\phi+1);\ \phi_2 = -\phi \Rightarrow |\phi_1+\phi_2| = 1 \Rightarrow$ NON-STATIONARY!

ARIMA(1, 1, 1) in Random Shock Form

- $(1-\phi B)(1-B)y_t = (1-\theta B)a_t$

  compared with random shock form: $y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \cdots$

  $(1-\phi B)(1-B)(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \cdots) = (1-\theta B)a_t$

  $[1-(\phi+1)B+\phi B^2](1+\psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \cdots) = (1-\theta B)$

  $1 + [\psi_1 - (\phi+1)]B + [\psi_2 - \psi_1(\phi+1) + \phi]B^2 + [\psi_3 - \psi_2(\phi+1) + \psi_1\phi]B^3 + \cdots$
  $+ \sum_{j=4}^{\infty}[\psi_j - \psi_{j-1}(\phi+1) + \psi_{j-2}\phi]B^j = (1-\theta B)$

  $$\begin{cases} \psi_1 - (\phi+1) = -\theta \\ \psi_2 - \psi_1(\phi+1) + \phi = 0 \Rightarrow \psi_2 = \psi_1(\phi+1) - \phi \\ \vdots \\ \psi_j - \psi_{j-1}(\phi+1) + \psi_{j-2}\phi = 0, \forall j \geq 3 \end{cases}$$

ARIMA(1, 1, 1) in Inverted Form

- $(1-\phi B)(1-B)y_t = (1-\theta B)a_t$

  $(1-\theta B)^{-1}(1-\phi B)(1-B)y_t = a_t$

  $\left(1+\theta B+\theta^2 B^2+\cdots\right)\left[1-(\phi+1)B+\phi B^2\right]y_t = a_t$

  $$\left\{1+[\theta-(\phi+1)]B+\left[\theta^2-\theta(\phi+1)+\phi\right]B^2+\sum_{j=3}^{\infty}\left[\theta^j-\theta^{j-1}(\phi+1)+\theta^{j-2}\phi\right]B^j\right\}y_t = a_t$$

  compared with inverted form: $y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \pi_3 y_{t-3} + \cdots + a_t$.

  - The key requirement: $|\theta| < 1$, will, at the same time, ensure $\sum_{i=1}^{\infty}|\pi_i| < \infty$, and $\lim_{j\to\infty}|\pi_i| = 0$.

An Invertible but Non-stationary ARIMA(p, d, q) has

The special property: $\sum_{i=1}^{\infty} \pi_i = 1.$

Proof:

$\because \phi_p(B)(1-B)^d y_t = \theta_q(B) a_t$ is invertible and non-stationary,

$\therefore d \geq 1$ and the roots of $\theta_q(B)$ are outside the unit circle.

$\Rightarrow a_t = \pi(B) y_t = \left(1 - \sum_{i=1}^{\infty} \pi_i B^i\right) y_t$ satisfies $\phi_p(B)(1-B)^d = \theta_q(B)\pi(B)$

Let $B = 1 \Rightarrow \phi_p(1)(1-1)^d = \theta_q(1)\pi(1) = 0.$

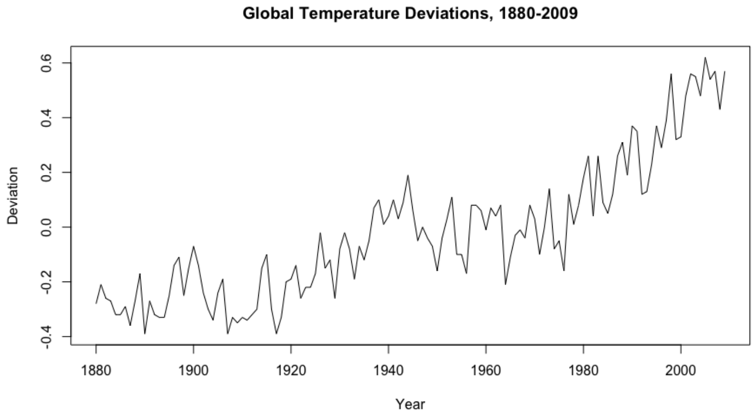
However, $\theta_q(1) \neq 0 \Rightarrow \pi(1) = \left(1 - \sum_{i=1}^{\infty} \pi_i 1^i\right) = 0 \Rightarrow \sum_{i=1}^{\infty} \pi_i = 1.$

ACF of a Non-stationary Time Series

- Given a stationary ARMA(p, q),

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \cdots + \phi_p\rho_{k-p},$$
$$\forall k > q$$

$$\text{As } \phi(B) = \prod_{i=1}^{p}(1 - G_i B),$$



  - where $G_i^{-1}, i = 1, \ldots p$, are the roots to $\phi(B) = 0$.
  - Box & Jenkins (1976) proved: $\rho_k = A_1G_1^k + A_2G_2^k + \cdots + A_pG_p^k$, A_i are constants. To make the series stationary, $|B_i| > 1 \Rightarrow |G_i| < 1$. If $|G_i|$ is far from 1^-, ρ_k will vanish quickly.
    - In the opposite, if one or more $|G_i| \to 1^-$, ACF doesn't change a lot with time.

Some Special Non-stationary Time Series Processes

2

Random Walk

- Given an AR(1) model: $y_t = \phi y_{t-1} + a_t$,
  - what if $\phi = 1$? $y_t = y_{t-1} + a_t \Rightarrow \nabla y_t = a_t \Rightarrow (1-B)y_t = a_t$
  - the random shock form: $y_t = \psi(B)a_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$

$$(1-B)\psi(B)a_t = a_t \Rightarrow (1-B)\left(1 + \psi_1 B + \psi_2 B^2 + \cdots\right) = 1$$

$$1 + (\psi_1 - 1)B + (\psi_2 - \psi_1)B^2 + \sum_{j=3}^{\infty}(\psi_j - \psi_{j-1})B^j = 1 \Rightarrow \psi_1 = 1 = \psi_2 = \psi_3 = \cdots \psi_j$$

$$y_t = a_t + a_{t-1} + a_{t-2} + \cdots + a_1$$

  - If the level L is not zero, i.e., the series is not centered,

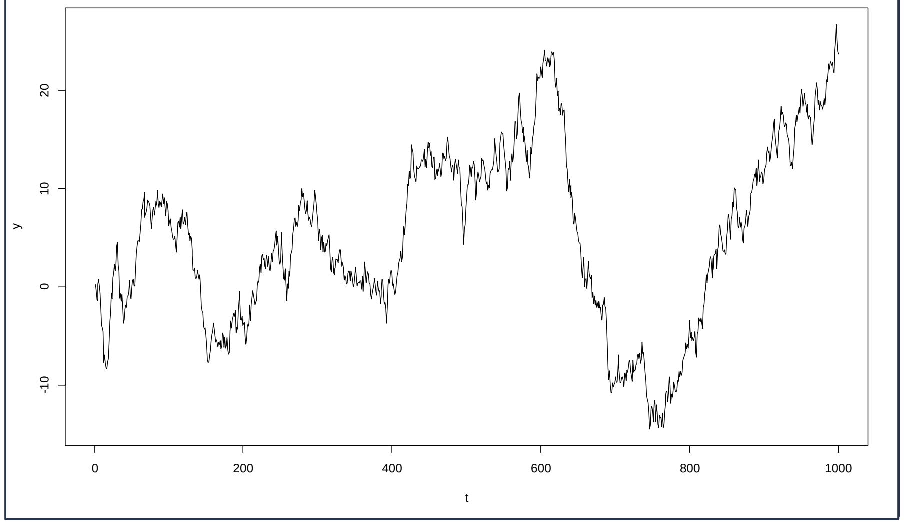
$$y_t = tL + \sum_{i=1}^{t} a_i$$

Random Walk Example

$\nabla y_t = L + a_t;$

$L = 0;$

$a_t \sim \mathrm{N}(0, 1^2).$

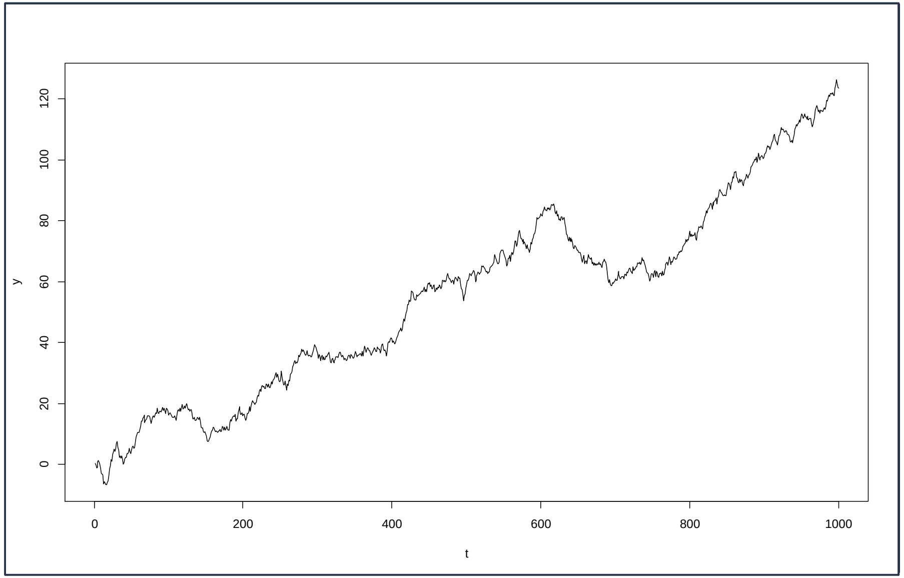


Random Walk Example

$\nabla y_t = L + a_t;$

$L = 0.1;$

$a_t \sim \mathrm{N}(0, 1^2).$



IMA(1, 1)

- $(1-B)y_t = (1-\theta B)a_t$
  - the random shock form: $y_t = \psi(B)a_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$

$$(1-B)\psi(B)a_t = (1-\theta B)a_t \Rightarrow (1-B)\left(1+\psi_1 B + \psi_2 B^2 + \cdots\right) = (1-\theta B)$$

$$1 + (\psi_1 - 1)B + (\psi_2 - \psi_1)B^2 + \sum_{j=3}^{\infty}(\psi_j - \psi_{j-1})B^j = (1-\theta B)$$

$$\psi_1 - 1 = -\theta \Rightarrow \psi_1 = 1-\theta = \psi_2 = \psi_3 = \cdots = \psi_j$$

$$y_t = a_t + (1-\theta)a_{t-1} + (1-\theta)a_{t-2} + \cdots + (1-\theta)a_1$$

  - $\sum_i^{\infty} \psi_i^2 \to \infty \Rightarrow y_t$ is non-stationary.

IMA(1, 1) in Inverted Form

- $(1-B)y_t = (1-\theta B)a_t = (1-\theta B)\left(1-\pi_1 B-\pi_2 B^2-\cdots\right)y_t$

$$\Rightarrow (1-B) = (1-\theta B)\left(1-\pi_1 B-\pi_2 B^2-\cdots\right)$$

$$= 1-(\theta+\pi_1)B-(\pi_2-\pi_1\theta)B^2-(\pi_3-\pi_2\theta)B^3-\sum_{j=4}^{\infty}\left(\pi_j-\pi_{j-1}\theta\right)B^4$$

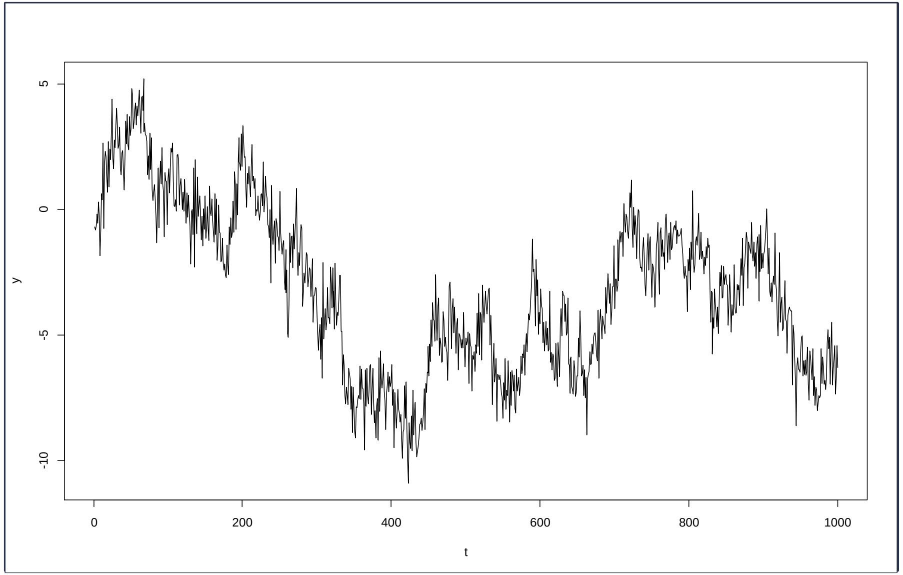
$$\begin{cases} \theta+\pi_1=1 \\ \pi_2-\pi_1\theta=0 \\ \pi_j-\pi_{j-1}\theta=0 \end{cases} \Rightarrow \begin{cases} \pi_1 = 1-\theta \\ \pi_2 = \pi_1\theta = \theta(1-\theta) \\ \vdots \\ \pi_j = \pi_{j-1}\theta = \pi_{j-2}\theta^2 = \cdots = \theta^{j-1}(1-\theta) \end{cases}$$

$$y_t = (1-\theta)y_{t-1}+\theta(1-\theta)y_{t-2}+\cdots+\theta^{j-1}(1-\theta)y_{t-j}+\cdots+a_t$$

Exponential Smoothing Model

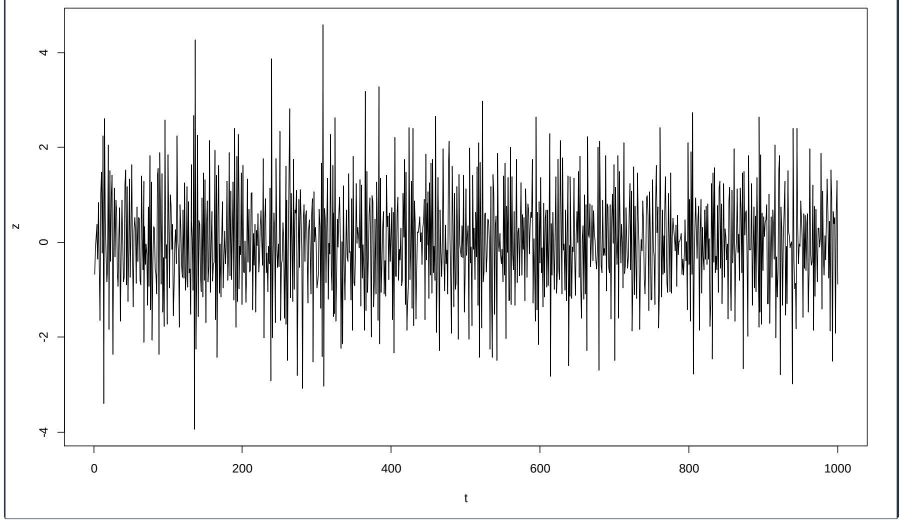
IMA (1, 1)

$$y_t = y_{t-1} + a_t - 0.6a_{t-1}.$$



IMA (1, 1)

$\nabla y_t = a_t - 0.6a_{t-1}$.



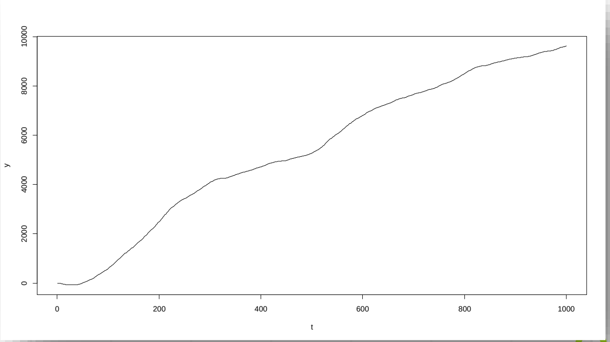
IMA(2, 0)

$$\nabla^2 y_t = (1-B)^2 y_t = a_t$$

- the random shock form: $y_t = \psi(B)a_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$

$$(1-B)^2(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots) = a_t$$
$$(1-2B+B^2)(1+\psi_1 B + \psi_2 B^2 + \cdots) = 1$$
$$1 + (\psi_1 - 2)B + (\psi_2 - 2\psi_1 + 1)B^2 + (\psi_3 - 2\psi_2 + \psi_1)B^3 + \cdots + (\psi_j - 2\psi_{j-1} + \psi_{j-2})B^j = 1$$

$$\begin{cases} \psi_1 = 2 \\ \psi_2 = 2\psi_1 - 1 = 3 \\ \psi_3 = 2\psi_2 - \psi_1 = 4 \\ \vdots \\ \psi_j = 2\psi_{j-1} - \psi_{j-2} = j+1 \end{cases}$$



$$y_t = a_t + 2a_{t-1} + 3a_{t-2} + 4a_{t-3} + \cdots = \sum_j^{\infty} (j+1)a_{t-j}$$

Let $z_t = y_t - y_{t-1}$, the series becomes $\nabla z_t = a_t$, a random walk.

Summary of Non-stationary Time Series

- Transformation is applied: $\nabla y_t = y_t - y_{t-1}$; $\sqrt{y_t}$; $\ln(y_t)$; ...
- When d in ∇^d is too large, the series loses its meaning.
- Homogeneous Non-stationary model: $\text{ARIMA}(p,d,q) \Rightarrow \phi(B)\nabla^d y_t = \theta(B)a_t + L$.
- Three explicit forms
  - Differencing Equation Form
    $$y_t = \phi_1 y_{t-1} + \cdots + \phi_{p+d} y_{t-p-d} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$
  - Random Shock Form
    $$y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \cdots$$
  - Inverted Form
    $$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \pi_3 y_{t-3} + \cdots + a_t$$
- ACF of $\text{ARIMA}(p,d,q)$ is difficult to observe.