Analysis of Time Series, L2

Rolf Larsson

Uppsala University

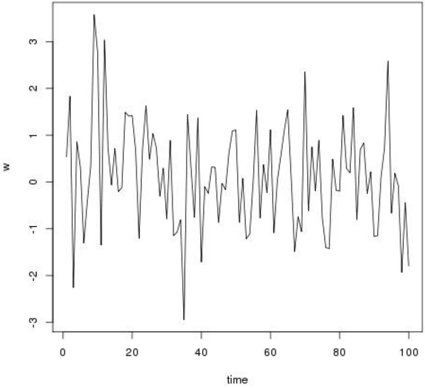
26 mars 2025

Today

- 1.5: Stationary time series
- 1.6: Estimation of correlation
- 2.3: Differencing
- Menti

Stationary time series

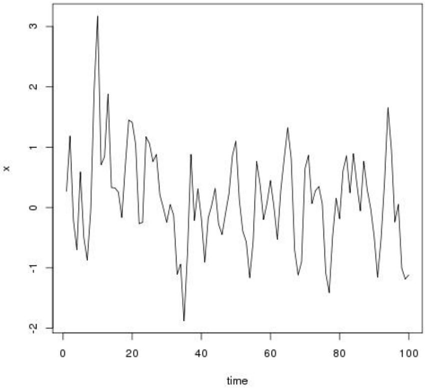
Example 1: White noise, $w_t \sim N(0, \sigma_w^2)$, independent



Erratic behaviour.

Stationary time series

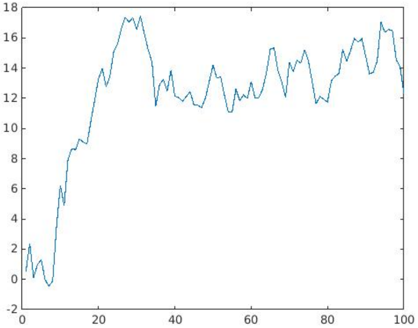
Example 2: Moving average, $x_t = \frac{1}{2}(w_t + w_{t-1})$



A little smoother.

Stationary time series

Example 3: Random walk, $x_t = x_{t-1} + w_t$



Recall: $E(x_t) = \mu_t = 0$,
$\mathrm{cov}(x_s, x_t) = \gamma(s, t) = \min(s, t)\sigma_w^2 \Rightarrow \mathrm{var}(x_t) = \gamma(t, t) = t\sigma_w^2$.

Stationary time series

Definition (1.6)

If

$$\{x_{t_1}, x_{t_2}, \ldots, x_{t_k}\} \quad \text{and} \quad \{x_{t_1+h}, x_{t_2+h}, \ldots, x_{t_k+h}\}$$

have identical joint distributions for all choices of $(t_1, t_2, \ldots, t_k)$ and h, then $\{x_t\}$ is *strictly stationary*.

Stationary time series

Special cases:

- $k = 1$: x_t and x_{t+h} have identical distributions for all t and h.
- Implications:
  - The mean function μ_t is constant.
  - The variance function $\gamma(t, t)$ is constant.
- $k = 2$: (x_s, x_t) and (x_{s+h}, x_{t+h}) have identical joint distributions for all (s, t) and h.
- Implications:
  - The autocovariance function $\gamma(s, t)$ only depends on s and t through $|s - t|$.
  - The autocorrelation function $\rho(s, t)$ only depends on s and t through $|s - t|$.

Stationary time series

Definition (1.7)

If x_t has finite variance for all t,

1. the mean function μ_t is constant,
2. the autocovariance function $\gamma(s, t)$ only depends on s and t through $|s - t|$,

then $\{x_t\}$ is *weakly stationary*.

- $\mathrm{var}(x_t) < \infty$ and x_t strictly stationary $\Rightarrow x_t$ is weakly stationary.
- $\Leftarrow$ not true in general. (cf problem 1.16)

Stationary time series

Definition (1.13)

A process $\{x_t\}$ is said to be *Gaussian* if $\{x_{t_1}, x_{t_2}, ..., x_{t_k}\}$ has a multivariate normal distribution for all choices of $(t_1, t_2, ..., t_k)$.

For a Gaussian process, the concepts of strict and weak stationarity are equivalent.

Stationary time series

Definition (1.8)

The *autocovariance function* of a *stationary* stochastic process $\{x_t\}$ is defined as

$$\gamma(h) = \mathrm{cov}(x_{t+h}, x_t).$$

Definition (1.9)

The *autocorrelation function* of a *stationary* stochastic process $\{x_t\}$ is defined as

$$\rho(h) = \mathrm{corr}(x_{t+h}, x_t) = \frac{\gamma(h)}{\gamma(0)}.$$

Stationary time series

Calculate $\gamma(h)$ and $\rho(h)$ for

1. the white noise process w_t.
2. the moving average process $x_t = \frac{1}{2}(w_t + w_{t-1})$.

Stationary time series

Some properties:

1. $\mathrm{var}(x_t) = \gamma(0)$
2. $\gamma(h) = \gamma(-h)$ for all h
3. $\rho(h) = \rho(-h)$ for all h
4. $|\rho(h)| \leq 1$ for all h
5. $|\gamma(h)| \leq \gamma(0)$ for all h

Stationary time series

Recall:

Definition (1.4)

The *cross-covariance function* between two series $\{x_t\}$ and $\{y_t\}$ is defined as

$$\gamma_{xy}(s, t) = \mathrm{cov}(x_s, y_t).$$

Definition (1.5)

The *cross-correlation function* between two series $\{x_t\}$ and $\{y_t\}$ is defined as

$$\rho_{xy}(s, t) = \mathrm{corr}(x_s, y_t) = \frac{\gamma_{xy}(s, t)}{\sqrt{\gamma_x(s, s)\gamma_y(t, t)}}.$$

Stationary time series

Definition (1.10)

Two stationary time series $\{x_t\}$ and $\{y_t\}$ are said to be *jointly (weakly) stationary* if the cross-covariance function

$$\gamma_{xy}(h) = \mathrm{cov}(x_{t+h}, y_t)$$

is a function only of h.

Definition (1.11)

The *cross-correlation function* between two jointly stationary time series $\{x_t\}$ and $\{y_t\}$ is defined as

$$\rho_{xy}(h) = \mathrm{corr}(x_{t+h}, y_t) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}}.$$

Stationary time series

- Let $x_t = \frac{1}{2}(w_t + w_{t-1})$ and $y_t = w_t$.
- Calculate $\gamma_{xy}(h)$ and $\rho_{xy}(h)$.

Stationary time series

Some properties:

1. $|\rho_{xy}(h)| \leq 1$ for all h
2. $\rho_{xy}(h) = \rho_{yx}(-h)$ for all h
3. $\rho_{xy}(h)$ and $\rho_{xy}(-h)$ are not equal in general!

Estimation of correlation

Definition (1.14)

The sample autocovariance function is defined as

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}),$$

for $h = 0, 1, ..., n-1$, $\hat{\gamma}(-h) = \hat{\gamma}(h)$.

Definition (1.15)

The sample autocorrelation function (ACF) is defined as

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

Estimation of correlation

Theorem (Property 1.2)

If the process $\{x_t\}$ is white noise with finite fourth moment, then for large n,

$$\hat{\rho}(h) \approx N\left(0, \frac{1}{n}\right).$$

Stricter formulation: $\sqrt{n}\hat{\rho}(h)$ converges to $N(0,1)$ as $n \to \infty$.

Proof: see theorem A.7.

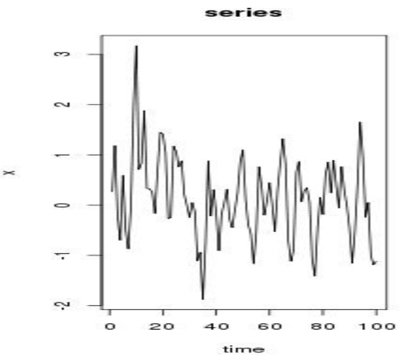
- For large n, H_0: $\rho(h) = 0$ is rejected vs H_1: $\rho(h) \neq 0$ at approximately the 5% level if

$$|\hat{\rho}(h)| > \frac{2}{\sqrt{n}}.$$

- In R: `acf(x)`

Estimation of correlation

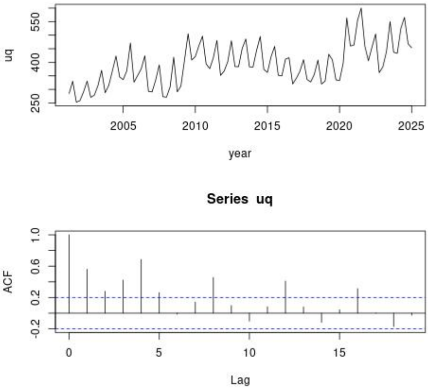
Simulation of $x_t = \frac{1}{2}(w_t + w_{t-1})$



The ACF is significantly nonzero only for lag 1.

Estimation of correlation

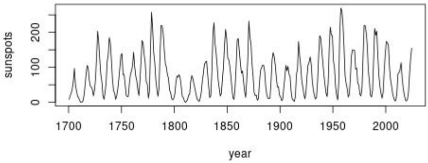
Unemployment, quarterly data.

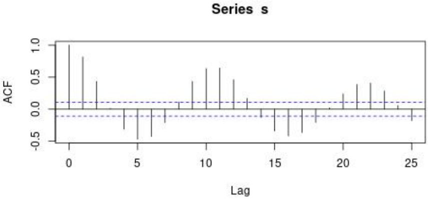


The ACF is large for multiples of 4.

Estimation of correlation

Average number of sunspots per year.

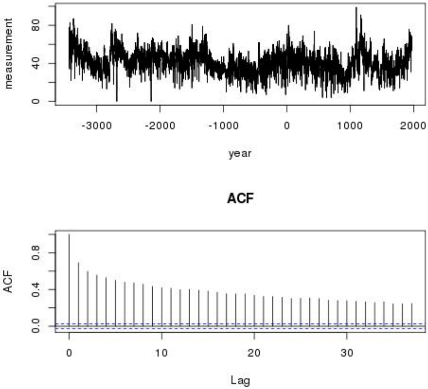




The ACF is large for multiples of around 10 or 11.

Estimation of correlation

Tree ring measurements, Mount Campito.



The ACF decays slowly with increasing lag. ('Long memory'.)

Estimation of correlation

Definition (1.16)

The sample cross-covariance function is defined as

$$\hat{\gamma}_{xy}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(y_t - \bar{y}),$$

for $h = 0, 1, ..., n-1$, $\hat{\gamma}_{xy}(-h) = \hat{\gamma}_{yx}(h)$.

The sample cross-correlation function (CCF) is defined as

$$\hat{\rho}_{xy}(h) = \frac{\hat{\gamma}_{xy}(h)}{\sqrt{\hat{\gamma}_x(0)\hat{\gamma}_y(0)}}.$$

Estimation of correlation

Theorem (Property 1.3)

If $\{x_t\}$ and $\{y_t\}$ are independent and at least one of $\{x_t\}$ or $\{y_t\}$ is independent white noise, then for n large,

$$\hat{\rho}_{xy}(h) \approx N\left(0, \frac{1}{n}\right).$$

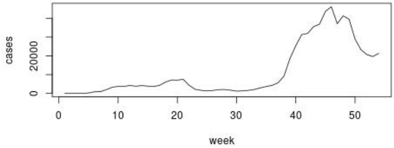
Stricter formulation: $\sqrt{n}\hat{\rho}_{xy}(h)$ converges to $N(0,1)$ as $n \to \infty$.

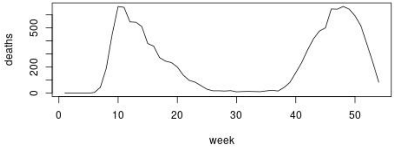
Proof: see theorem A.8.

In R: `ccf(x,y)`

Estimation of correlation

Covid-19, numbers of reported cases and deaths in Sweden, 2020-2021 week 6.

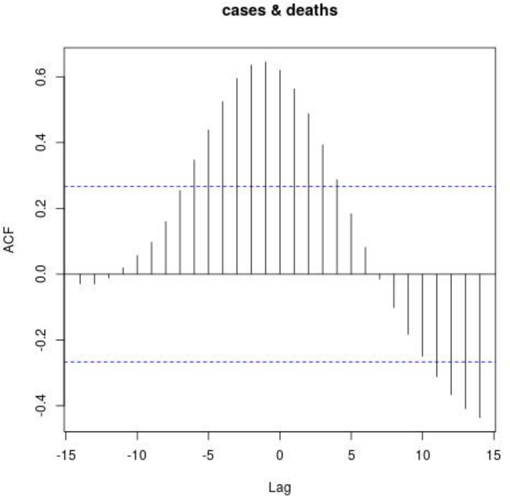




Not much testing in the beginning.

Estimation of correlation

Estimated cross correlation, cases and deaths. Largest at lag -1!



Differencing

Definition (2.4)

The *backshift operator* is defined by

$$Bx_t = x_{t-1}.$$

For $k = 1, 2, \ldots$, $B^k x_t = x_{t-k}$.

Definition (2.5)

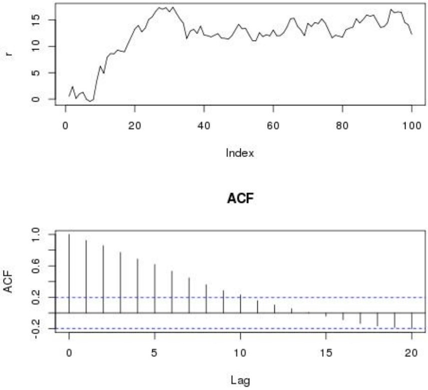
Differences of order d are defined by

$$\nabla^d x_t = (1 - B)^d x_t.$$

- Special cases:
  - $\nabla^1 x_t = \nabla x_t = (1 - B)x_t = x_t - x_{t-1}$ (often needed)
  - $\nabla^2 x_t = (1 - B)^2 x_t = x_t - 2x_{t-1} + x_{t-2}$ (rarely needed)
- It is *very rare* that more than two differences are needed!
- In R: `diff(x)`

Differencing

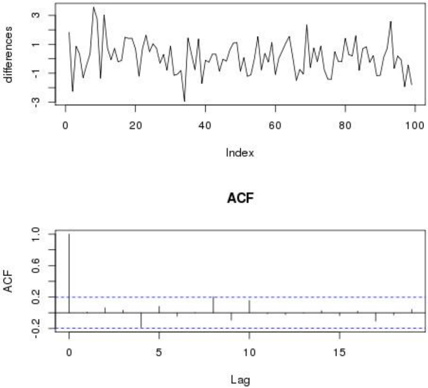
Simulated random walk.



The ACF decays slowly for increasing lags.

Differencing

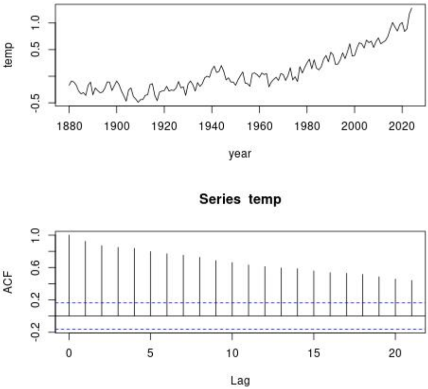
Simulated random walk, differences.



The ACF cuts off after lag zero. It is between the dashed lines for pos. lags.

Differencing

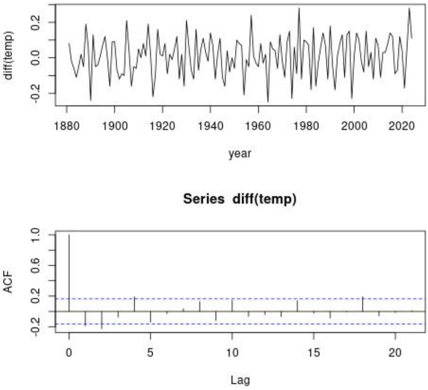
Global mean temperature.



Increasing trend. The ACF decays slowly for increasing lags.

Differencing

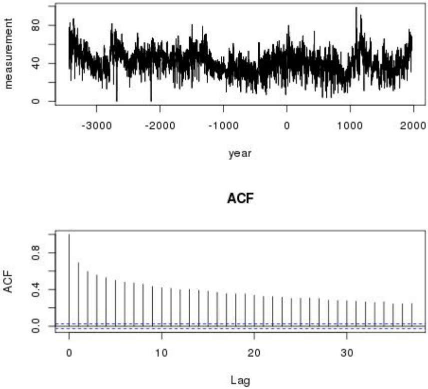
Global mean temperature, differences.



The ACF basically cuts off after lag zero.

Differencing

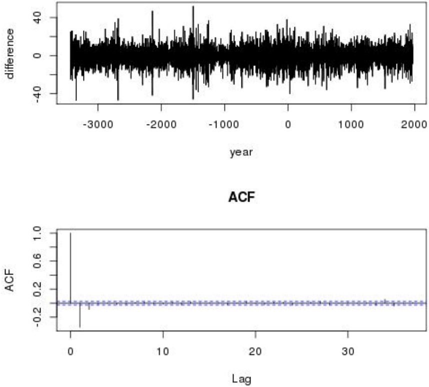
Tree ring measurements, Mount Campito.



The ACF decays slowly with increasing lag. ('Long memory'.)

Differencing

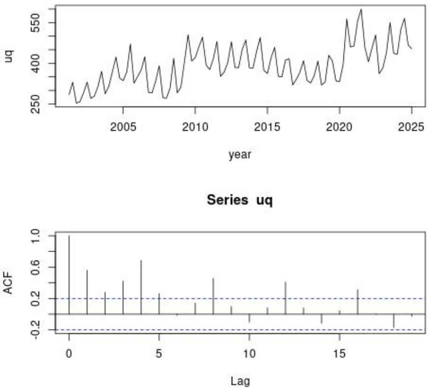
Tree ring measurements, Mount Campito, differences.



The ACF is markedly different from zero only for small lags.

Differencing

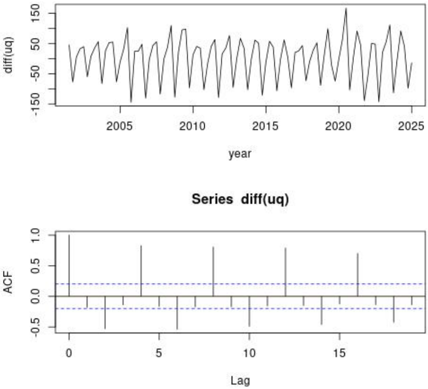
Unemployment (quarterly).



The ACF is large for multiples of 4.

Differencing

Unemployment, differences.



Still, the ACF is large for multiples of 4.

News of today

- strict stationarity
- weak stationarity
- ACF
- CCF
- differencing