3: Integrerede processer (ARIMA) og long-memory modeller (ARFIMA)

Tidsrækkeanalyse

Kasper Rosenkrands

Ikke stationær tidsrække

Hvis vi har at gøre med data som ikke er stationært, siger vi at vores tidsrække er integreret. Til dette hører en orden, forstået på den måde at hvis vores tidsrække er integreret af orden et betyder det at vi opnår en stationær tidsrække ved at differense en gang.

Et eksempel på en ikke stationær tidsrække er en **random walk**.

En random walk er en AR(1) process med $\phi = 1$

$$x_t = x_{t-1} + w_t,$$

med andre ord er det en sum af hvidstøjsled.

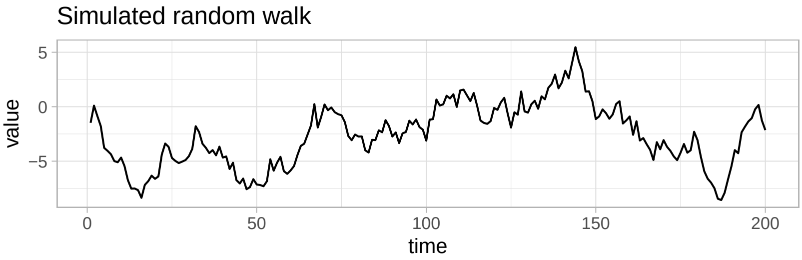
Før jeg givet et eksempel på en random walk vil jeg introducere **differensoperatoren**.

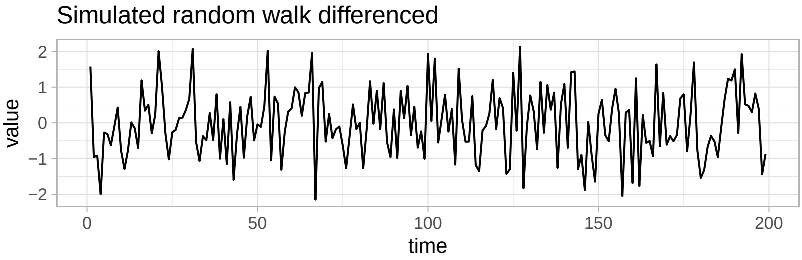
Differens operatoren

Den almindelige differensoperator er givet ved

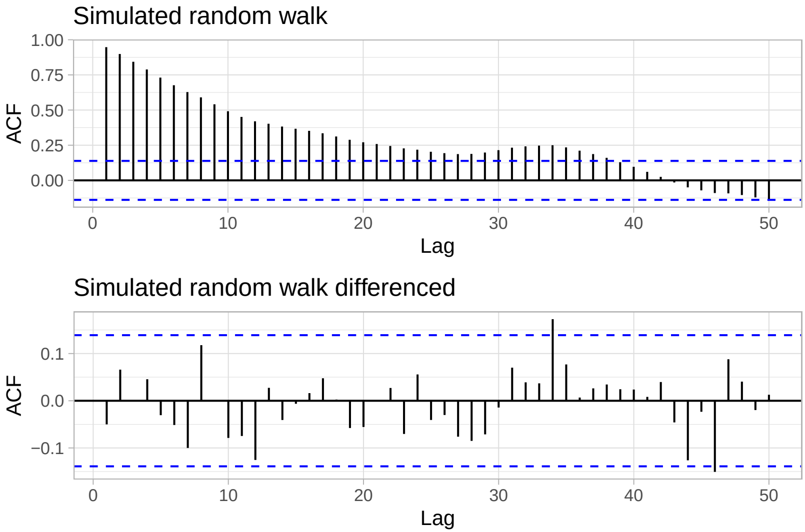
$$\nabla x_t = (1 - B)x_t = x_t - x_{t-1}.$$

Eksempel på random walk





Eksempel på random walk ACF



ARIMA model

Generelt siger vi, for $p, d, q \geq 0$, at tidsrækken $\{x_t\}_{t \in \mathbb{Z}}$ er en ARIMA(p, d, q) proces hvis

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

er en ARMA(p,q). I så fald skriver vi at

$$\phi(B) \nabla^d x_t = \theta(B) w_t.$$

Long memory modeller

Vi indeler tidsrækker i to kategorier

1. Kort hukommelse

$$\sum |\gamma(h)| < \infty.$$

2. Lang hukommelse

$$\sum |\gamma(h)| = \infty.$$

For at behandle tidsrækker med lang hukommelse benyttes **ARFIMA**.

ARFIMA model

For ikke-negative heltal p, q samt $-0.5 < d < 0.5$, siger vi at tidsrækken $\{x_t\}_{t\in\mathbb{Z}}$ er en **ARFIMA**(p, d, q) proces hvis

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

er ARMA(p, q).

Forskellen mellem ARIMA og ARIFMA er at d i ARIMA er begrænset til et heltal, hvor vi i ARFIMA ikke har denne begrænsing.

Den normale differens operator giver dog ikke længere mening når d antager heltals værdier, derfor introducere vi fraktionel differens:

$$\nabla^d = (1 - B)^d = 1 + \sum_{j=1}^{\infty} \frac{\Gamma(j - d)}{\Gamma(-d)j!} B^j$$

hvor gamma funktionen er defineret ved

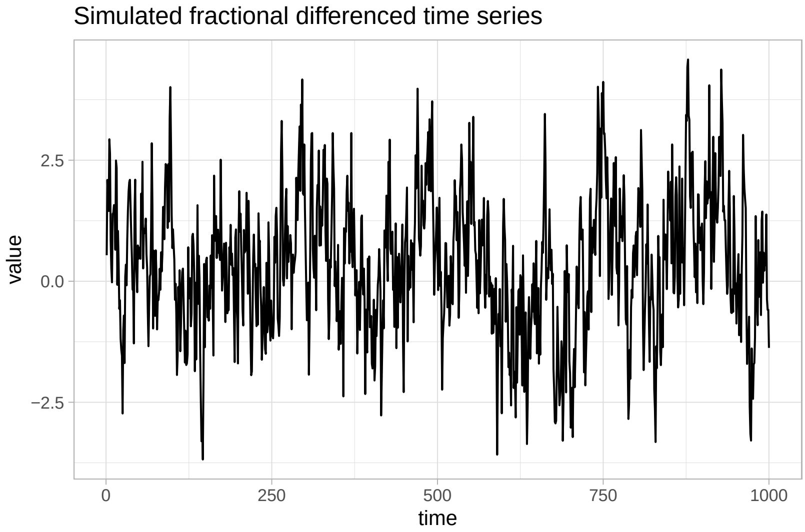
$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \text{ for } x \neq 0, -1, -2, \dots$$

Fraktionel differens eksempel

For at give et eksempel på en lang hukommelse vil jeg simulere en sådan tidsrække

```
library(fracdiff)
set.seed('123')
y <- fracdiff.sim(
  1000,
  ar = .6,
  ma = .25,
  d = .3
)
```

Plot af den simulerede tidsrække



Informations kriterier

```r
AIC(fit11 <- fracdiff(y$series, nar = 1, nma = 1))
```

```
## [1] 2828.756
```

```r
AIC(fit10 <- fracdiff(y$series, nar = 1, nma = 0))
```

```
## [1] 2830.518
```

```r
AIC(fit01 <- fracdiff(y$series, nar = 0, nma = 1))
```

```
## [1] 2834.78
```

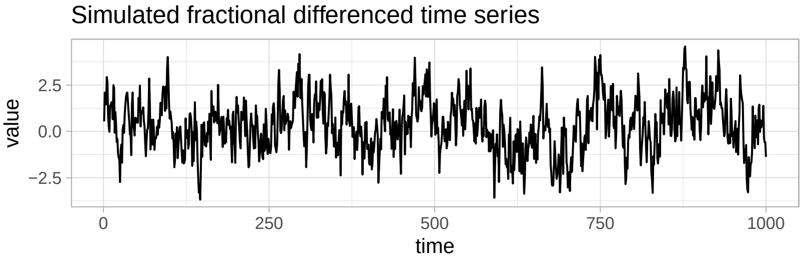
```r
AIC(fit00 <- fracdiff(y$series, nar = 0, nma = 0))
```

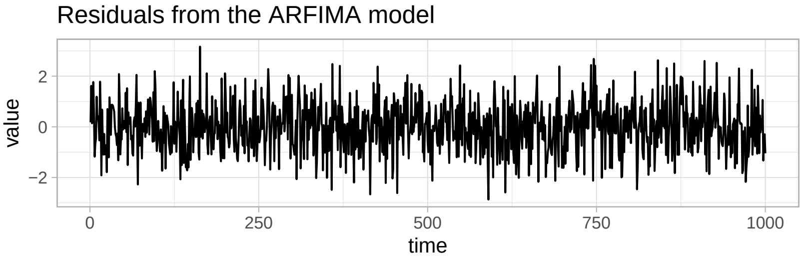
```
## [1] 2849.879
```

Koefficienter

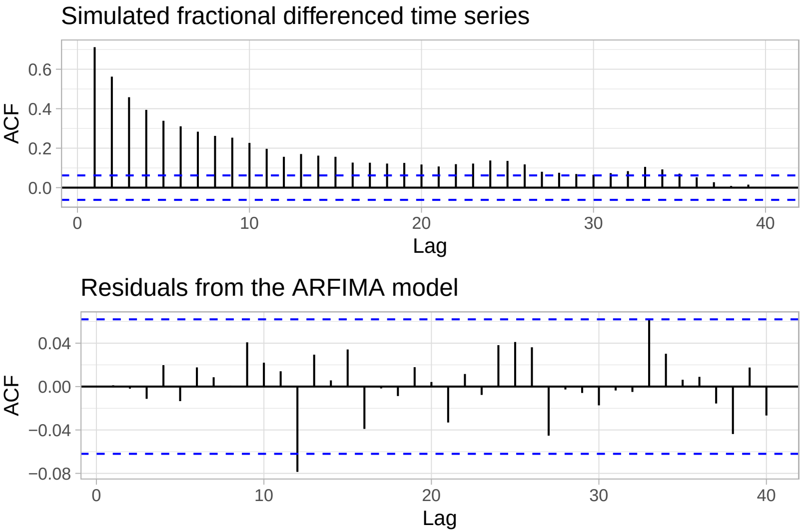
```r
fit11$d
```

```
## [1] 0.2808727
```





ACF plot


Simulated fractional differenced time series
ACF
0.6
0.4
0.2
0.0
0
10
20
30
40
Lag
Residuals from the ARFIMA model
ACF
0.04
0.00
−0.04
−0.08
0
10
20
30
40
Lag