The basics of time series data

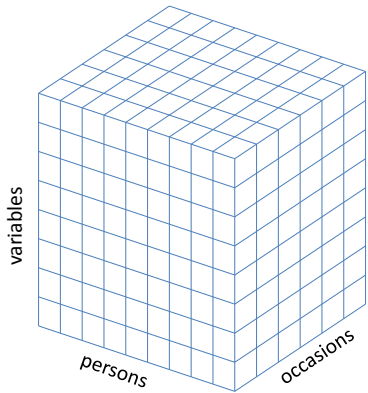
Modeling Intensive Longitudinal Data

Ellen Hamaker

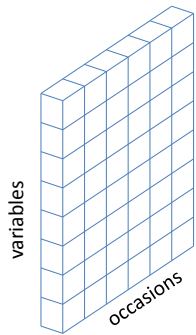
Utrecht University

Sharing science, *shaping tomorrow*

Cattell's data box

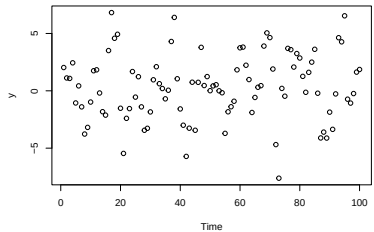


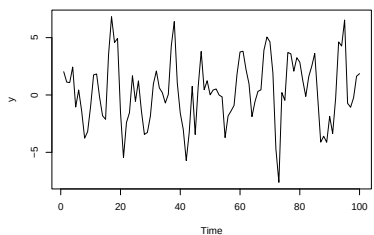
Time series data (N=1 ILD)



Univariate time series

Time	y
1	2.03
2	1.12
3	1.08
4	2.44
5	-1.06
6	0.43
7	-1.40
8	-3.77
9	-3.19
10	-0.98
11	1.75
12	1.83
13	-0.19
...	...
97	-1.07
98	-0.24
99	1.63
100	1.86



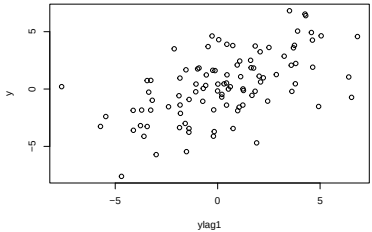


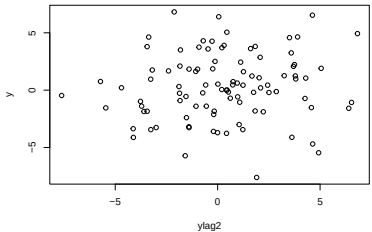
Utrecht University

Lagged variables

Time	y	y lag 1	y lag 2
1	2.03		
2	1.12	2.03	
3	1.08	1.12	2.03
4	2.44	1.08	1.12
5	-1.06	2.44	1.08
6	0.43	-1.06	2.44
7	-1.40	0.43	-1.06
8	-3.77	-1.40	0.43
9	-3.19	-3.77	-1.40
10	-0.98	-3.19	-3.77
11	1.75	-0.98	-3.19
12	1.83	1.75	-0.98
13	-0.19	1.83	1.75
...	...	...	...
97	-1.07	-0.73	6.54
98	-0.24	-1.07	-0.73
99	1.63	-0.24	-1.07
100	1.86	1.63	-0.24
101		1.86	1.63
102			1.86

Lagged relations





Autocorrelation at lag k

Autocorrelation at lag k:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$

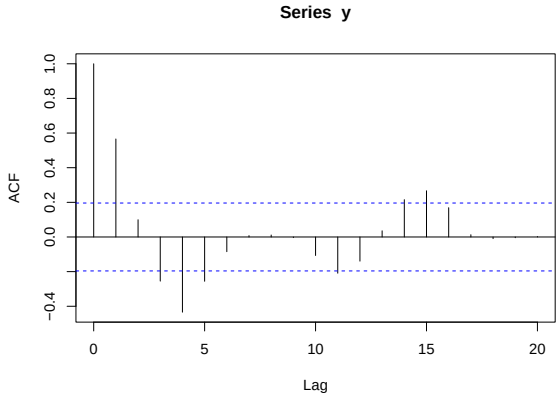
Variance (or: auto-covariance at lag 0):

$$\hat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^{T} \left(y_t - \bar{y}_t\right)^2$$

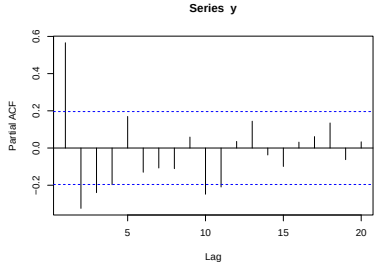
Auto-covariance at lag k:

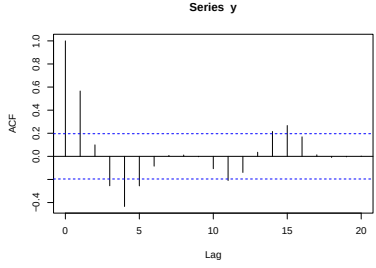
$$\hat{\gamma}_k = \frac{1}{T-k} \sum_{t=k+1}^{T} \left(y_t - \bar{y}_t\right)\left(y_{t-k} - \bar{y}_t\right)$$

Autocorrelation function (ACF)



Partial autocorrelation function (PACF)



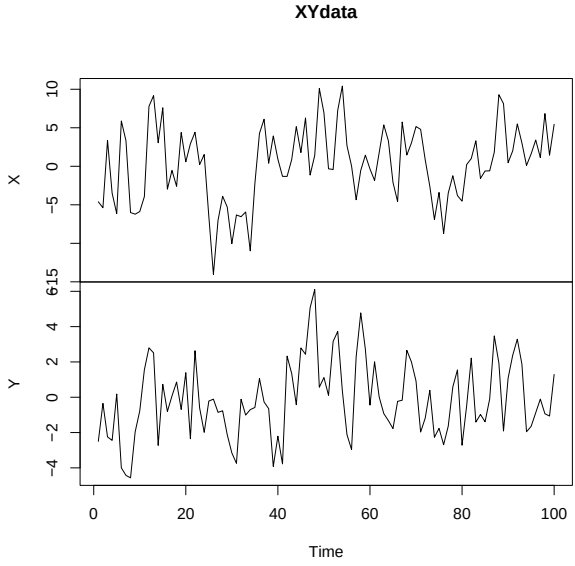


Time	y	y lag 1	y lag 2
1	2.03		
2	1.12	2.03	
3	1.08	1.12	2.03
4	2.44	1.08	1.12
5	-1.06	2.44	1.08
6	0.43	-1.06	2.44
7	-1.40	0.43	-1.06
8	-3.77	-1.40	0.43
9	-3.19	-3.77	-1.40
10	-0.98	-3.19	-3.77
11	1.75	-0.98	-3.19
12	1.83	1.75	-0.98
13	-0.19	1.83	1.75
...	...	...	...
97	-1.07	-0.73	6.54
98	-0.24	-1.07	-0.73
99	1.63	-0.24	-1.07
100	1.86	1.63	-0.24
101		1.86	1.63
102			1.86

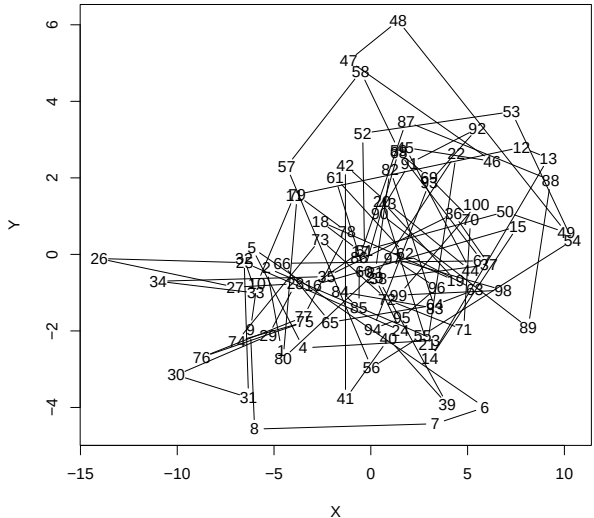
Multivariate data

Time	X	Y	X lag 1	Y lag 1
1	-4.61	-2.5	NA	NA
2	-5.38	-0.35	-4.61	-2.5
3	3.36	-2.24	-5.38	-0.35
4	-3.49	-2.44	3.36	-2.24
5	-6.15	0.18	-3.49	-2.44
…	…	…	…	…
99	1.44	-1.06	6.84	-0.94
100	5.45	1.29	1.44	-1.06

Visualize multivariate data: Sequence plots



Visualize data: State-space plot



Cross-correlations

Time	X	Y	X lag 1	Y lag 1
1	-4.61	-2.5	NA	NA
2	-5.38	-0.35	-4.61	-2.5
3	3.36	-2.24	-5.38	-0.35
4	-3.49	-2.44	3.36	-2.24
5	-6.15	0.18	-3.49	-2.44
...	...	...	...	...
99	1.44	-1.06	6.84	-0.94
100	5.45	1.29	1.44	-1.06

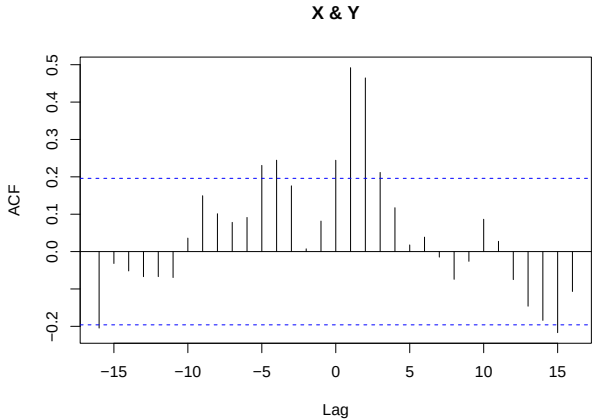
Correlation between X and Ylag1 (i.e., X_{t+1} and Y_t): 0.50

Correlation between Y and Xlag1 (i.e., X_{t-1} and Y_t): 0.10

Cross-correlation function (CCF)

Correlation between X_{t+1} and Y_t (lag 1): 0.50

Correlation between X_{t-1} and Y_t (lag -1): 0.10



Summary

- Time series data are characterized by the order of the observations
- This is the basis of making (univariate) sequence plots
- Time is included in state-space plots by connecting subsequent bivariate observations
- Dependencies over time are quantified with the autocorrelations and cross correlations

Utrecht University

Sharing science,
shaping tomorrow