Tues, 2-5-19
3. Autocorrelation

1. Lightning talk & Feedback on A2
2. Persistence
3. Autocorrelation function
4. Lagged scatterplot

Read notes_3.pdf

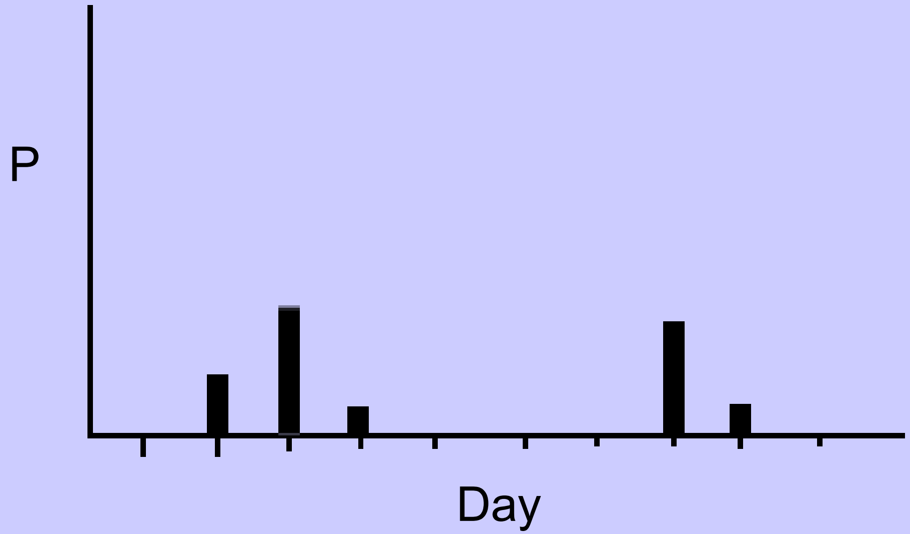
A2 Feedback

1. Download A2x.pdf from D2L
2. Automatic points, for running assignment and having uploaded by due time, is already marked in parentheses at top of first page
3. Each assignment has maximum possible 10 points; if you make no deductions, score is 10/10
4. Open your copy of the same assignment pdf you uploaded
5. In Acrobat Reader, using “Add text box,” mark in right margin for deductions only, with deduction and segment reference : (eg., -0.5 A)
6. At top of your pdf, mark grade like this : 9.5/10
7. If necessary, put any comments at top near the grade
8. Upload your self-graded pdf to folder A2_graded in D2L

Persistence

- "Tendency to stay in the same state or condition"
- Can be defined in terms of probability
- Reflects the dynamics of the system --- biological, physical, political, etc.

Example: daily precipitation



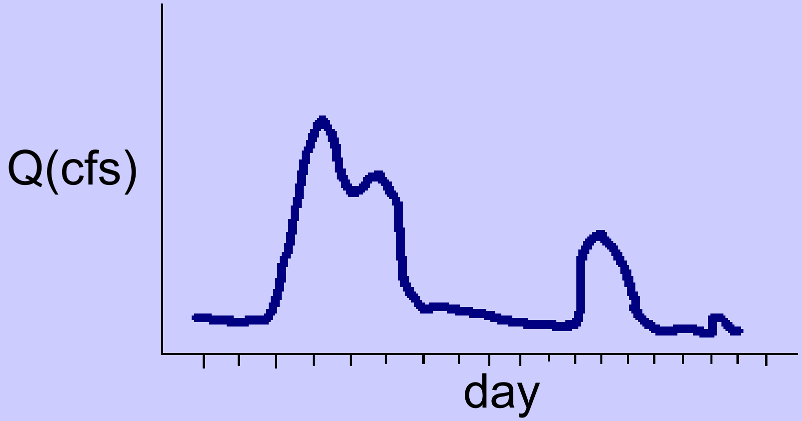
$$A_t = \text{event of precipitation on day } t$$

$$\Pr\left(A_t \mid A_{t-1}\right) > \Pr\left(A_t\right)$$

#events/#days

Driver: synoptic scale weather systems

Example: stream discharge



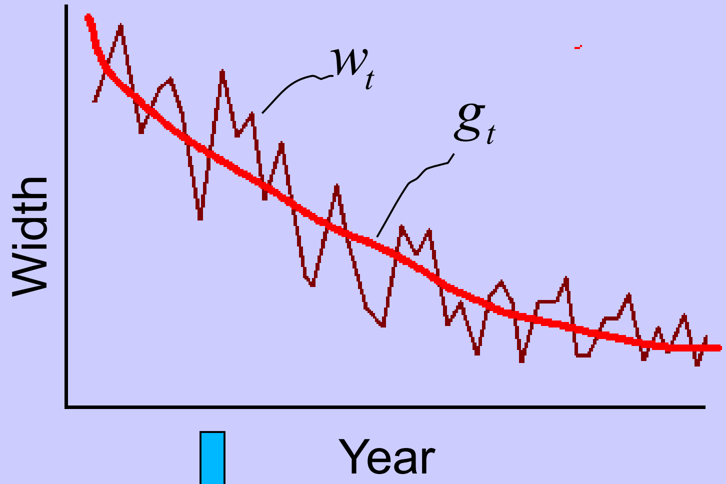
Sources of persistence

- Synoptic scale weather systems
- Travel time to gage
- Storage

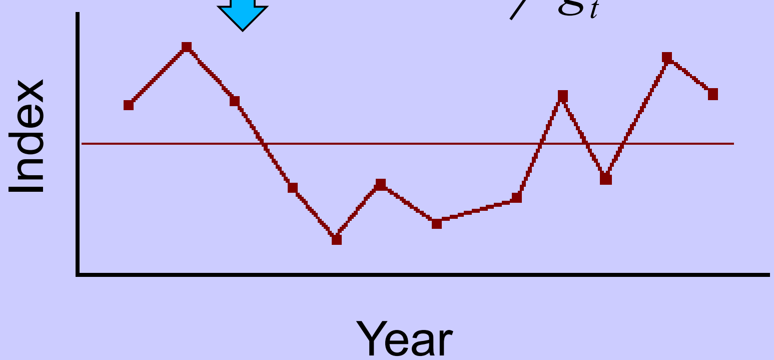
Sampling interval matters: daily flows might be highly persistent and annual flows ~ independent

Example: tree-ring width

- Ring width typically has trend, which is a form of persistence
- Trend removal gives an index, which still has persistence

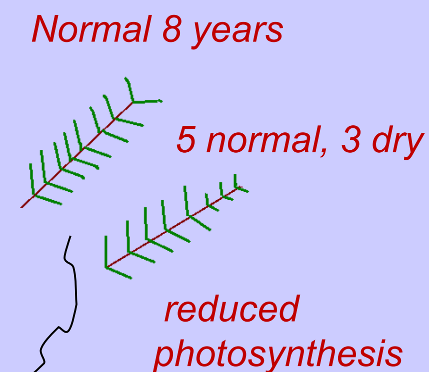


$$\text{Index} = \frac{w_t}{g_t}$$



Why?

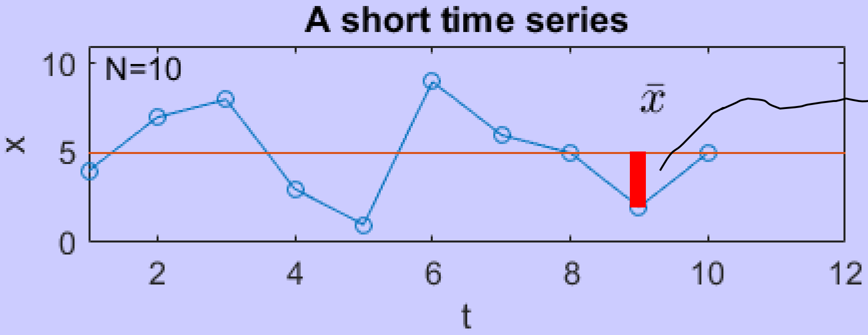
1. Food storage
2. Crown & root inertia
3. Needle retention
4. Soil moisture storage
5. Year-to-year persistence in climate



Autocorrelation

- Correlation with past values
- One particular mathematical expression of persistence

Autocorrelation function, sample (acf)



Departure from mean

Autocovariance at lag k

$$c(k) = \frac{1}{N} \sum_{t=1}^{N-k} \left(x_t - \overline{x}\right)\left(x_{t+k} - \overline{x}\right)$$

$$c(0) = s^2$$

<u>Not</u> (N-k)

Mean

$$\overline{x} = \frac{1}{N} \sum_{t=1}^{N} x_t$$

Variance

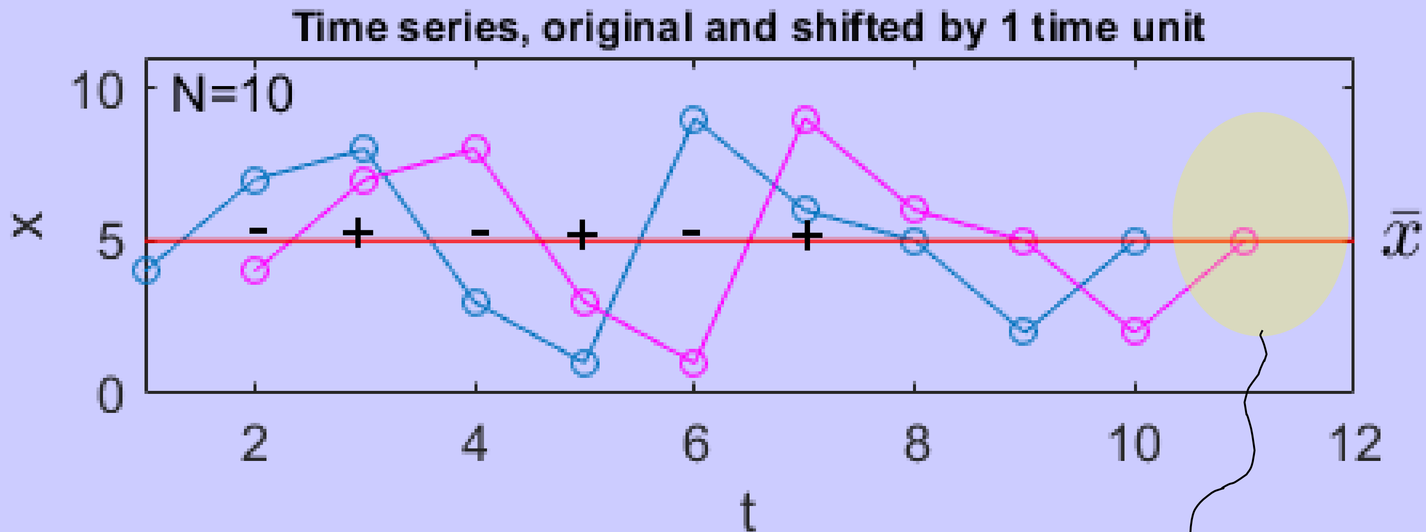
$$s^2 = \frac{1}{N} \sum_{t=1}^{N} \left(x_t - \overline{x}\right)^2$$

Autocorrelation at lag k

$$r(k) = \frac{c(k)}{c(0)}$$

$$-1 \le r \le 1, \quad r(0) \equiv 1$$

r(*k*) as correlation of series with Itself shifted *k* lags

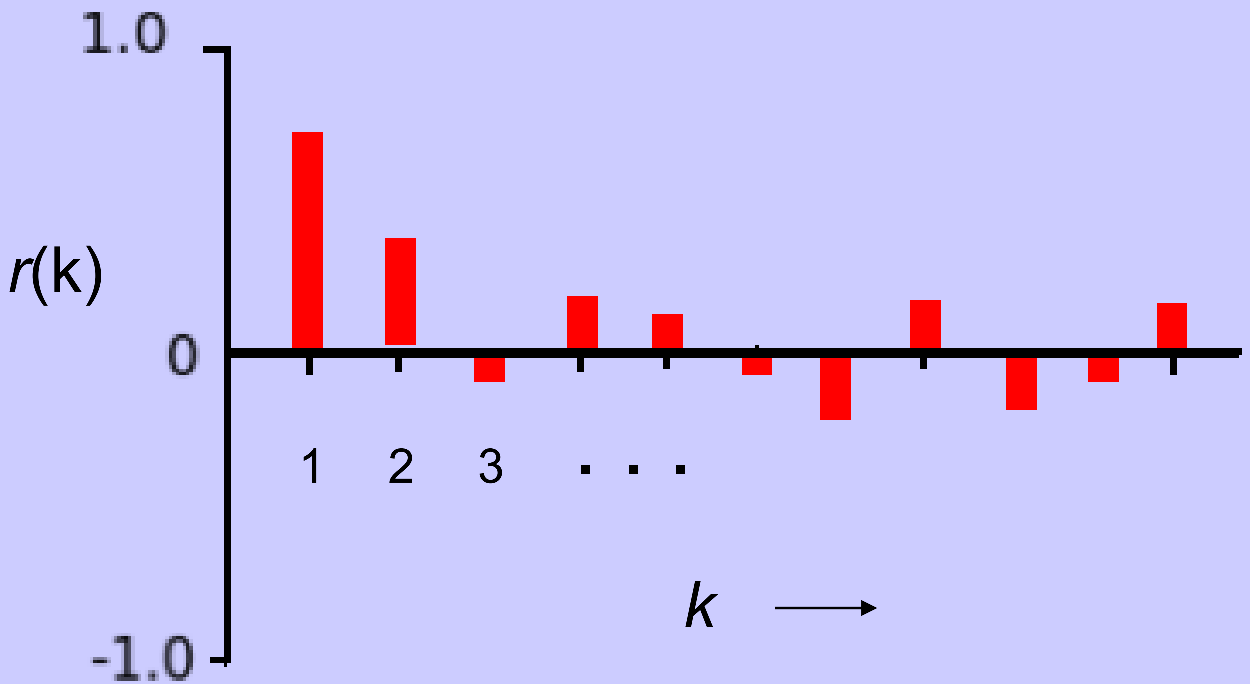


- Example here is for lag *k*=1
- Covariance is ~ average product of departures
- Sign of product marked above
- Acf is symmetric

K fewer data points for computing acf at each higher lag, k

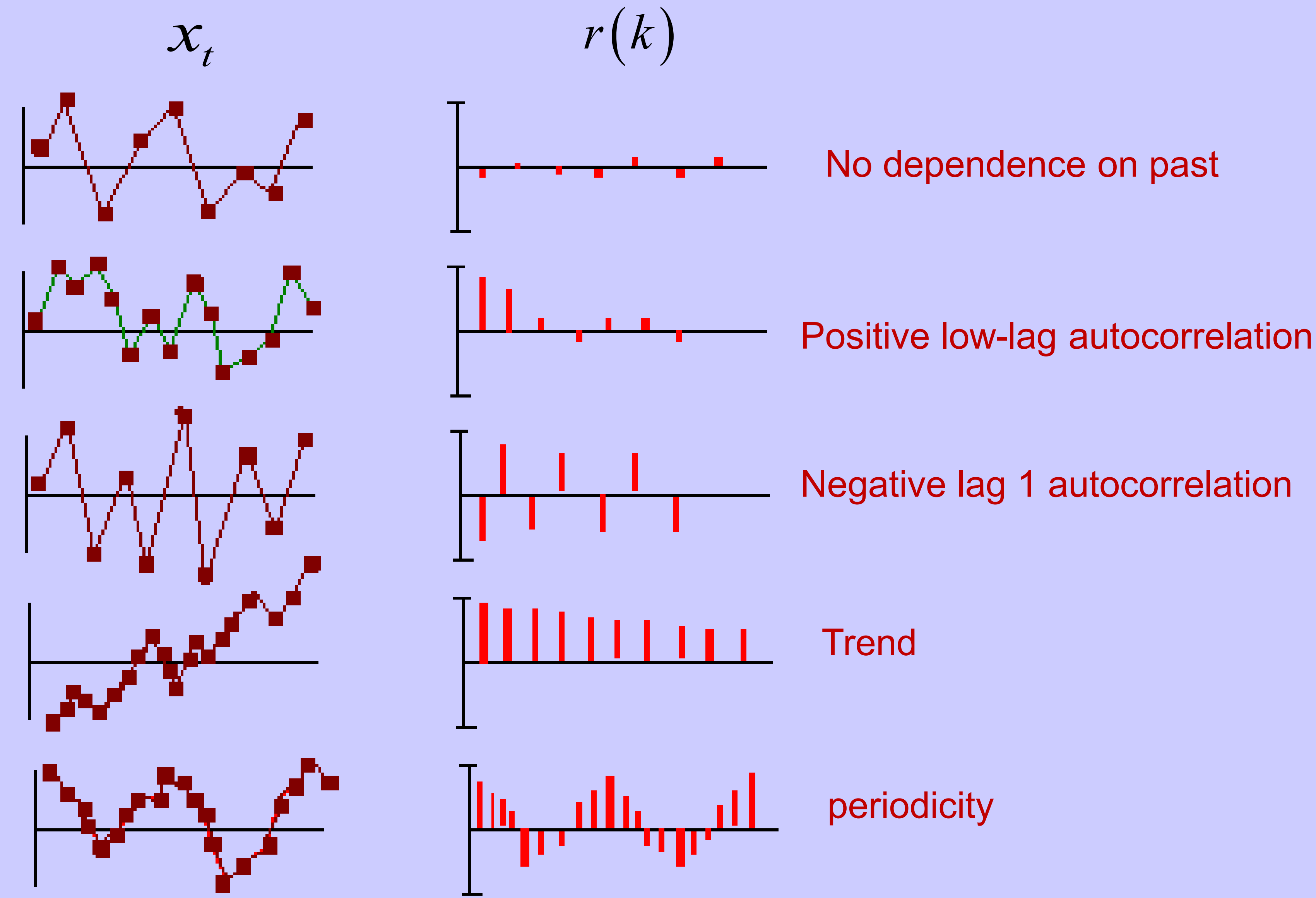
Computed r(1) could differ slightly from the Pearson r between the two plotted series because the mean used for r(1) is the global mean – based on all 10 observations

acf plot: *r*(*k*) plotted against lag, *k*



- Typically plotted out to lag 10-20
- Estimates increasingly uncertain as *k* becomes large fraction of *N*
- Characteristic patterns associated with certain patterns of persistence

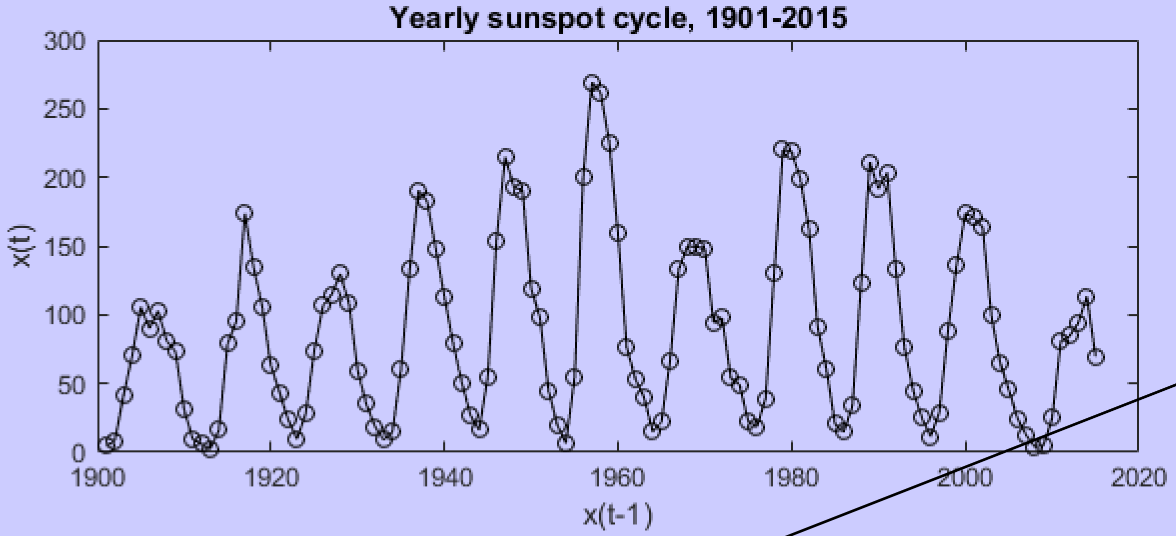
Characteristic acf patterns



Lagged scatterplot

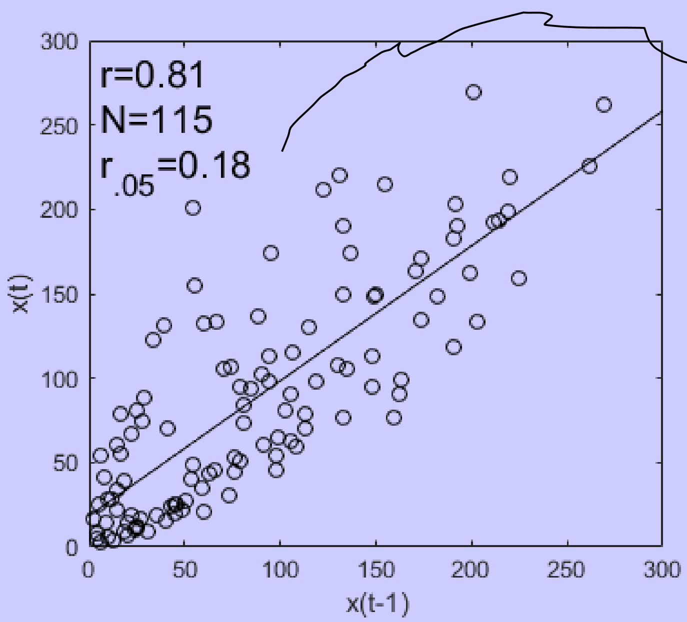
- The acf $r(k)$ for a specific lag k is approximately the same as the correlation coefficient of a time series with itself shifted by k time units
- A scatterplot of the time series against itself lagged k time units is another graphical way to check for autocorrelation at a specific lag.

Lagged scatterplot—sunspot example



Threshold correlation for 95% significance (2-tailed test

$$r_{0.05} \simeq \pm 1.96 / \sqrt{N}$$



- Scatterplot correlation is approximately autocorrelation r(1)
- Scatterplot gives much more information than a single r(1)
- Plot supports strong linear dependence of current year's sunspot number on last year's