health studies
Fitbit

Grades class attendance

Longitudinal Data and Time Series (4/17/2017)

Examples: Climate (CO_2 level, temperature, precip) | economic data (housing prices, stock (bond prices, interest rates)

Time Series: Measuring a single variable over points in time

Can be regularly or irregularly spaced time intervals:

Data structure

one person/object, J time points } X_1

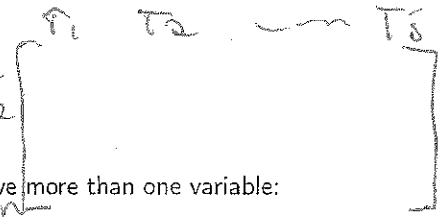
Regularly spaced (h=fixed #)

$T_i, T_{i+h}, T_{i+2h}, T_{i+3h}$

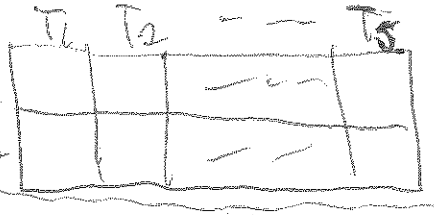
Irregularly spaced

$T_1, T_2, T_3 \ldots$

lots of people/objects, J time points } X_1, X_2, X_n



Can have more than one variable:



one person/object, two variables V_1, V_2

Several people/objects

several n x J matrices

Analyzing and Comparing Time Series

What are we interested in seeing with time series?

Trends { increasing? decreasing? no trend? } Cumulative measures (random increases or decreases)

Standard deviation
Range of possible values
Most common values

SPIKES (or valleys/dropoffs)

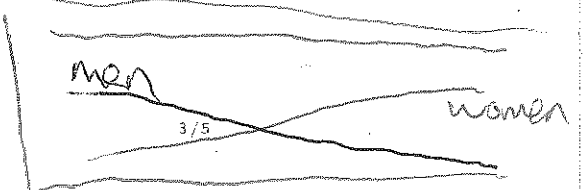
Periodicity: daily/weekly/yearly trends

parallel (↑↑ or ↓↓)

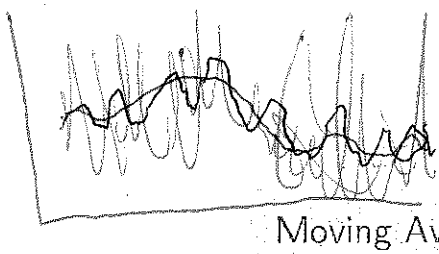
opposite (↑↓ or ↓↑)

What if we have multiple time series? How to compare?

* Must have measurements at the same time points!!!



Check for dependence on a subgroup of the data



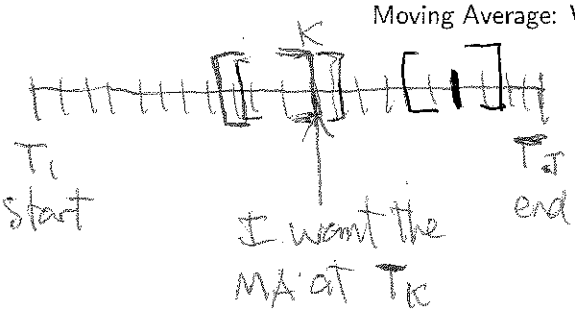
small ww
large ww
original TS

$ww = 1 \Rightarrow$ Not really a moving average – just gives back the original time series

Moving Average Plots

parameter that we get to choose

Moving Average: Want to visualize how the trend changes over time.



[ ] = "window"

$\longleftrightarrow$ = "window width" ↑ Big ww: smoother moving averages, global trends in TS

$MA_K = \frac{\sum_{i=K-ww+1}^{K} T_i}{ww}$

Small ww: jagged high-variation MAs, local trends in TS

Can downweight older observations in your moving window:

"weighted moving average" → often useful in forecasting / prediction

* How do we choose the weights? → set of parameters that you get to choose

↳ cross-validation

↳ weights $\propto 1/$["how far"]

4/5

Suppose $ww = 3$, you might choose something like weighting by "importance" Importance = $[5, 2, 1/2]$

$\Rightarrow$ weights = $\left[\frac{5}{5+2+1/2}, \frac{2}{5+2+1/2}, \frac{1/2}{5+2+1/2}\right]$ → scaled so that weights sum to 1

Lags and Autocorrelation

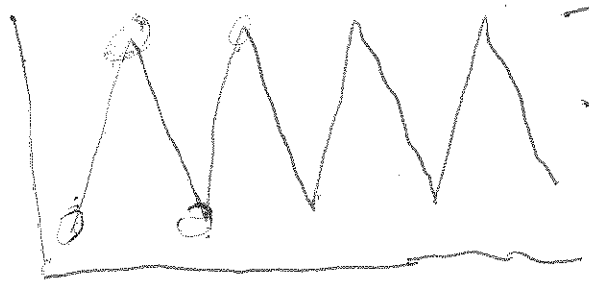
Lags: Does one time point influence future time point(s)?

Lag 1: looking at TS vs. itself one time step ago

Lag 2: ~~~~~ two ~~~~~

Lag K: ~~~~~ K ~~~~~

Autocorrelation: Correlation between a time series and a lagged version of itself



→ high negative AC at lag 1

→ high positive AC at lag 2

(high values at time K-2 correspond to high values at time K)

5/5

$$R(s,t) = E\left[(X_s - \mu_s)(X_t - \mu_t)\right]$$

↳ two time points

$\sigma_s \sigma_t$

Moving Average Plots

Moving Average: Want to visualize how the trend changes over time.

Can downweight older observations in your moving window:

Lags and Autocorrelation

Lags: Does one time point influence future time point(s)?

Autocorrelation: Correlation between a time series and a lagged version of itself