Introduction to Time Series[1]
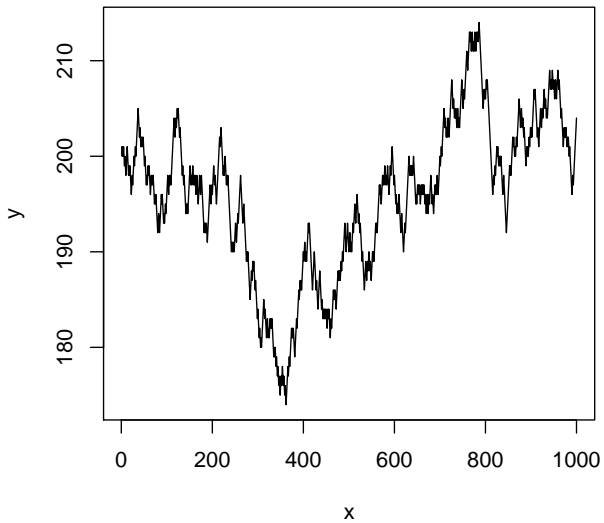
---

[1]This slide show is an open-source document. See last slide for copyright information.
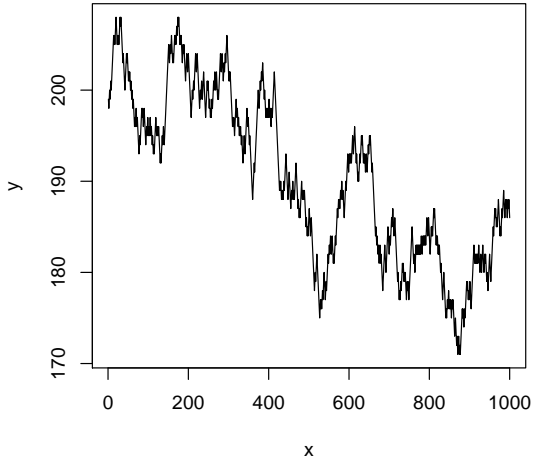
Time series

A sequence of measurements (random variables) $X_1, X_2, \ldots, X_n$

- Not a random sample.
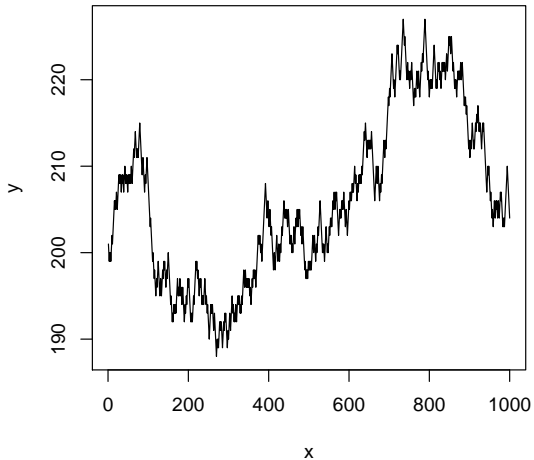- Not necessarily independent.
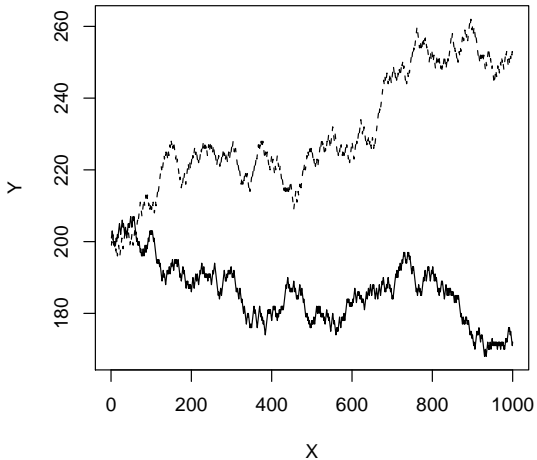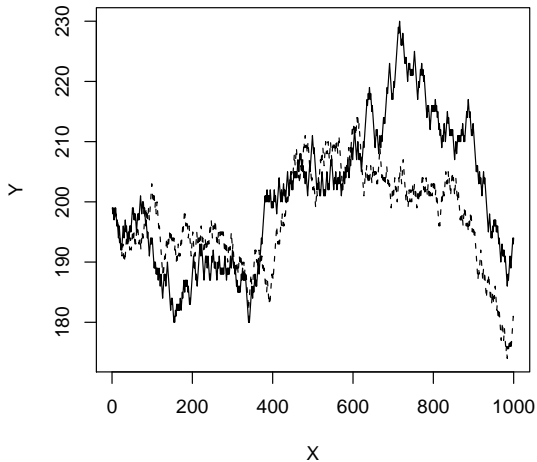- Sequentially dependent.

Trend, or Drift?

Trend, or Drift?

Trend, or Drift?

Related?

Related?

Correlations: 50 pairs of independent random walks, $n = 1000$ steps

Need around $|r| = 0.13$ for significance

```
-0.28175 -0.22242 -0.32170 -0.45053  0.07866  0.59167 -0.27414 -0.82570
-0.62175  0.43537  0.84147  0.04103 -0.17502 -0.89710 -0.19116 -0.53865
-0.50889  0.42855 -0.91074  0.90577  0.22818  0.84834 -0.52501  0.82583
-0.06838 -0.00234  0.16084  0.81393 -0.07063 -0.09908 -0.38405 -0.28510
 0.24850  0.12445  0.33509  0.33586  0.41241 -0.33482 -0.32021 -0.73808
 0.14045 -0.03618 -0.67757  0.81121 -0.39379 -0.58832 -0.26866  0.16687
 0.38541  0.12433
```

Random walk

Sometimes called Drunkard's walk

- Take a step left or right at random.
- Steps could be of variable length.
- Location at time t depends on location at time $t-1$.

$$X_t = X_{t-1} + \epsilon_t$$

$\epsilon_1, \epsilon_2, \ldots$ all independent and identically distributed.

Autoregressive Time Series

A generalization of the random walk

$$X_t = X_{t-1} + \epsilon_t \qquad \text{Random walk}$$

$$X_t = \beta_0 + \beta_1 X_{t-1} + \epsilon_t \qquad \text{First order autoregressive}$$

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \epsilon_t \qquad \text{Second order autogrgressive}$$

etc.

Stationary Time Series

- In a stationary time series, the distribution of X_t is not changing.
- In particular, all the X_t have the same mean and variance.

Expected value does not change

$$
\begin{aligned}
E(X_t) &= E(\beta_0 + \beta_1 X_{t-1} + \epsilon_t) \\
&= \beta_0 + \beta_1 E(X_{t-1}) + 0 \\
\Rightarrow \mu &= \beta_0 + \beta_1 \mu \\
\Rightarrow \beta_0 &= \mu(1 - \beta_1) \\
\Rightarrow \mu &= \frac{\beta_0}{1 - \beta_1}
\end{aligned}
$$

$$
\begin{aligned}
Var(X_t) &= Var(\beta_0 + \beta_1 X_{t-1} + \epsilon_t) \\
&= \beta_1^2 Var(X_{t-1}) + Var(\epsilon_t) \\
\Rightarrow \sigma^2 &= \beta_1^2 \sigma^2 + Var(\epsilon_t) \\
\Rightarrow Var(\epsilon_t) &= \sigma^2(1 - \beta_1^2)
\end{aligned}
$$

Covariance

$$
\begin{aligned}
Cov(X_{t-1}, X_t) &= Cov(X_{t-1}, \beta_0 + \beta_1 X_{t-1} + \epsilon_t) \\
&= \beta_1 Cov(X_{t-1}, X_{t-1}) + Cov(X_{t-1}, \epsilon_t) \\
&= \beta_1 Var(X_{t-1}) + 0 \\
&= \beta_1 \sigma^2
\end{aligned}
$$

So

$$
\begin{aligned}
Corr(X_{t-1}, X_t) &= \frac{\beta_1 \sigma^2}{\sqrt{\sigma^2}\sqrt{\sigma^2}} \\
&= \beta_1
\end{aligned}
$$

- The regression coefficient β_1 is usually denoted by ρ.
- The **First-order Autocorrelation**.
- Continuing the calculations, get $Corr(X_t, X_{t-2}) = \rho^2, \ldots$
- $Corr(X_t, X_{t-m}) = \rho^m$.
- So the correlation matrix looks like this:

$$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots \\ \rho & 1 & \rho & \rho^2 & \cdots \\ \rho^2 & \rho & 1 & \rho & \cdots \\ \rho^3 & \rho^2 & \rho & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

- Because $-1 < \rho < 1$, the pattern $\rho, \rho^2, \rho^3, \ldots$ displays a pattern of *exponential decay*: Graph it.
- Other time series structures have known signatures too.
- Higher-order autoregressive.
- Moving average.
- ARMA: Autoregressive Moving Average.
- Seasonal: Blips at seasonal lags.
- Non-stationary.
- Differencing is a big trick.
- ARIMA: Autoregressive Integrated Moving Average.
- Theorem: All the stationary processes can be approximated by autogregressive with enough lags.

Time series structures for the *error terms* (epsilons) in a regression

- What is the error term ϵ in a regression?
- Everything that affects y other than the x variables.
- Maybe those omitted variables are sequentially dependent.
- Like the temperature influences pop sales.
- Is sequential dependence likely? Depends on the logic of the data collection.
- Diagnose by the Durbin-Watson test and time series diagnostics on the residuals.

Durbin-Watson Test for Autocorrelation

- Usually, autocorrelation is positive.
- $H_0 : \rho = 0$ vs. $H_1 : \rho > 0$

$$D = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

- Reject when D is small. How small?
- Critical values and p-values are brutally hard to compute.
- Durbin and Watson published tables with upper and lower bounds for the critical values.
- Now SAS can compute all $p-$values, but it's an option.

What to do about autocorrelated residuals

- Try adding more explanatory variables, perhaps including time.
- Consider differencing.
- Directly model autocorrelated errors.

- Regression model with autogregressive errors: covers a lot of important cases.
- Especially in combination with *lagged* explanatory variables.
- Estimate β_j and ρ_k all at once by maximum likelihood.
- `proc autoreg` has many capabilities. As usual, we will explore just a few.
- Can you say GARCH?

Copyright Information

This slide show was prepared by Jerry Brunner, Department of Statistical Sciences, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. The LaTeX source code is available from the course website:

http://www.utstat.toronto.edu/brunner/oldclass/441s24