

# ARMA Models for Stationary Time Series

INSR 260, Spring 2009  
Bob Stine

# Overview

- ⌚ Stationarity

- ⌚ Correlation functions

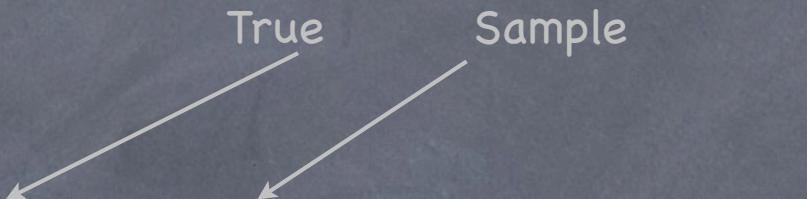
- ⌚ Autocorrelation function (TAC and SAC, ACF)
- ⌚ Partial autocorrelation function (TPAC and SPAC, PACF)

- ⌚ Models

- ⌚ Autoregression
- ⌚ Moving average
- ⌚ Combination

- ⌚ Examples

- ⌚ GDP, paper towel sales
- ⌚ Simulation of processes



# Stationarity

- ⦿ Stationarity is a key assumption in time series.
  - Stationarity = properties of the time series do not depend on when time starts.
    - Second-order stationary is most common form assumed in practice
  - Constant mean and variance
$$E(Y_t) = \mu \quad \text{Var}(Y_t) = \sigma_y^2$$
  - Autocorrelation only depends only on separation
$$\text{Corr}(Y_t, Y_{t-k}) = \rho_k$$
- ⦿ Why does stationarity matter?
  - Without stationarity, we've got nothing to average.
- ⦿ Two approaches to obtaining stationary data
  - Differencing                          Model changes  $Y_t - Y_{t-1}$
  - Regression

# Differencing

- Raw data has possible trends, but differencing removes most of this appearance
  - Visual procedure
    - Heuristic
    - Further differencing is also possible and sometimes needed
  - Later modeling allows more precise analysis

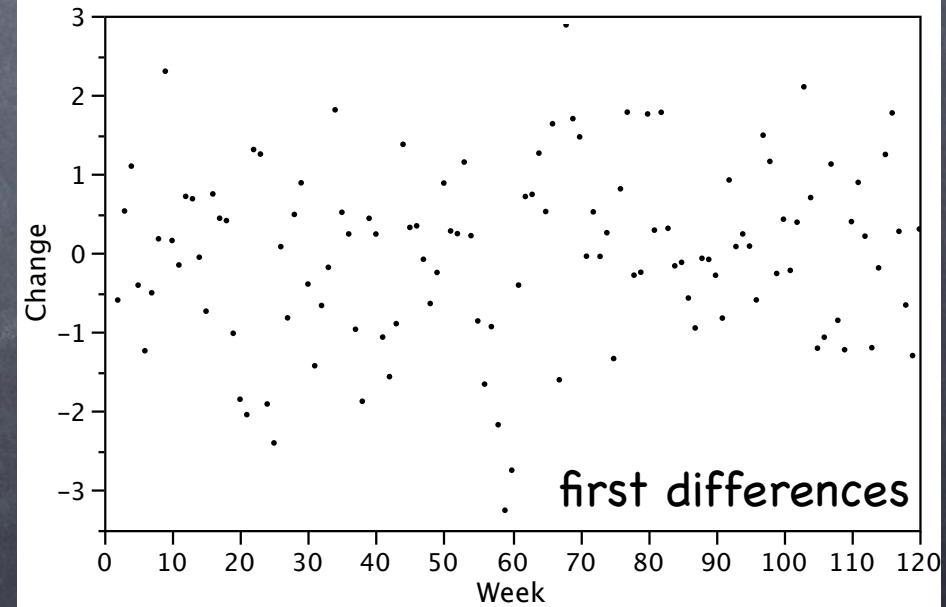
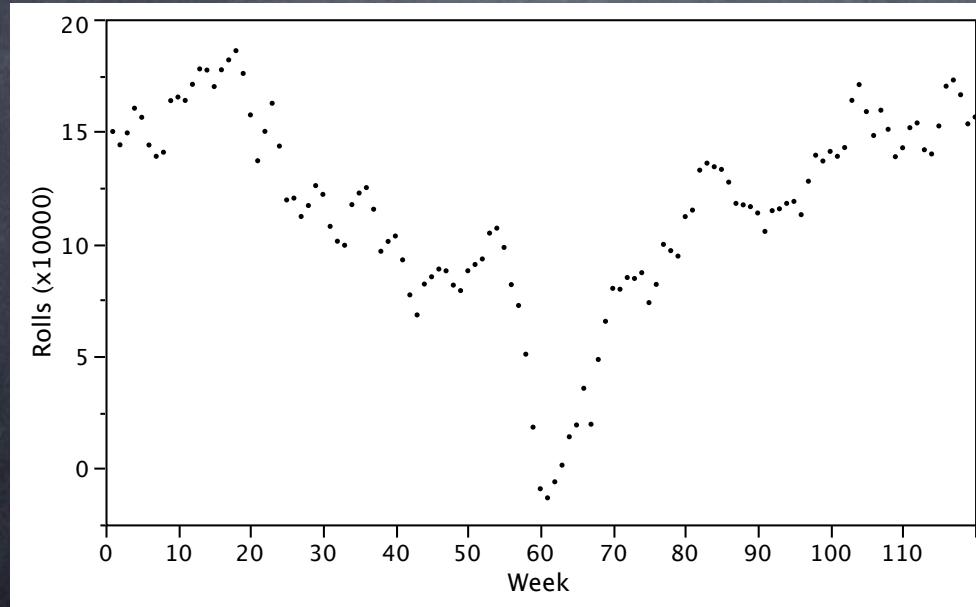
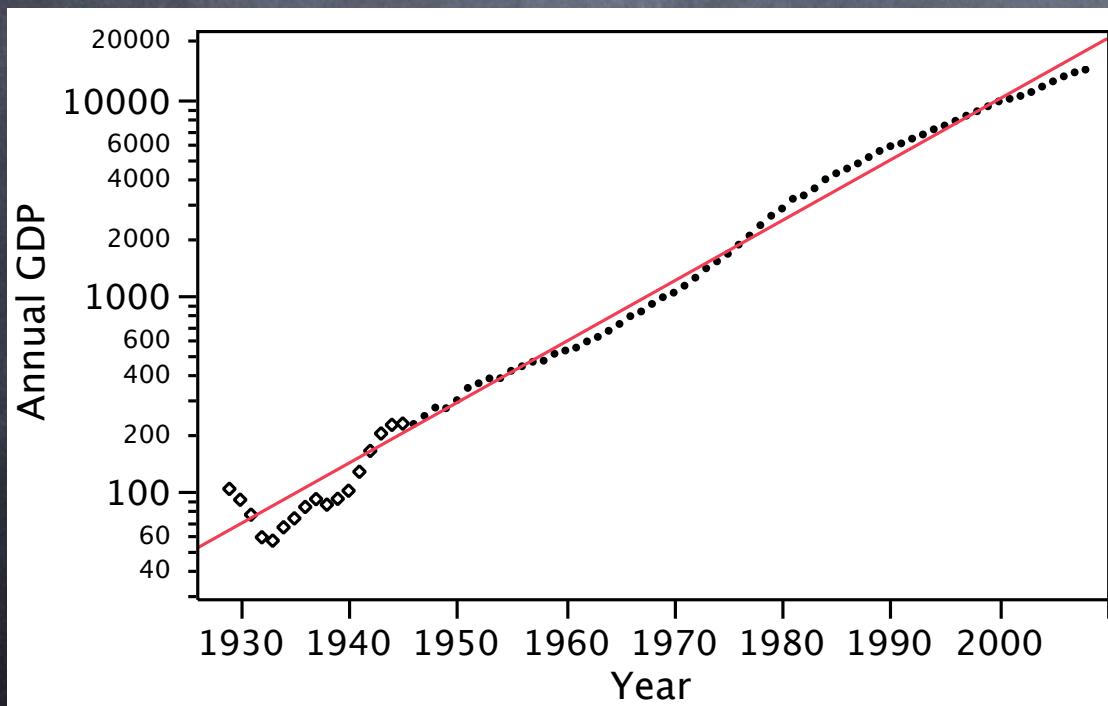


Table 9-1 Paper towels

# GDP Example

- Predict annual US GDP (real dollars)
  - Data span 1929–2008
  - Use data after WWII (1946–2008)
- Remarkably linear on log scale
  - $R^2 \approx 99\%$
  - About 7% annual rate of growth (includes inflation)

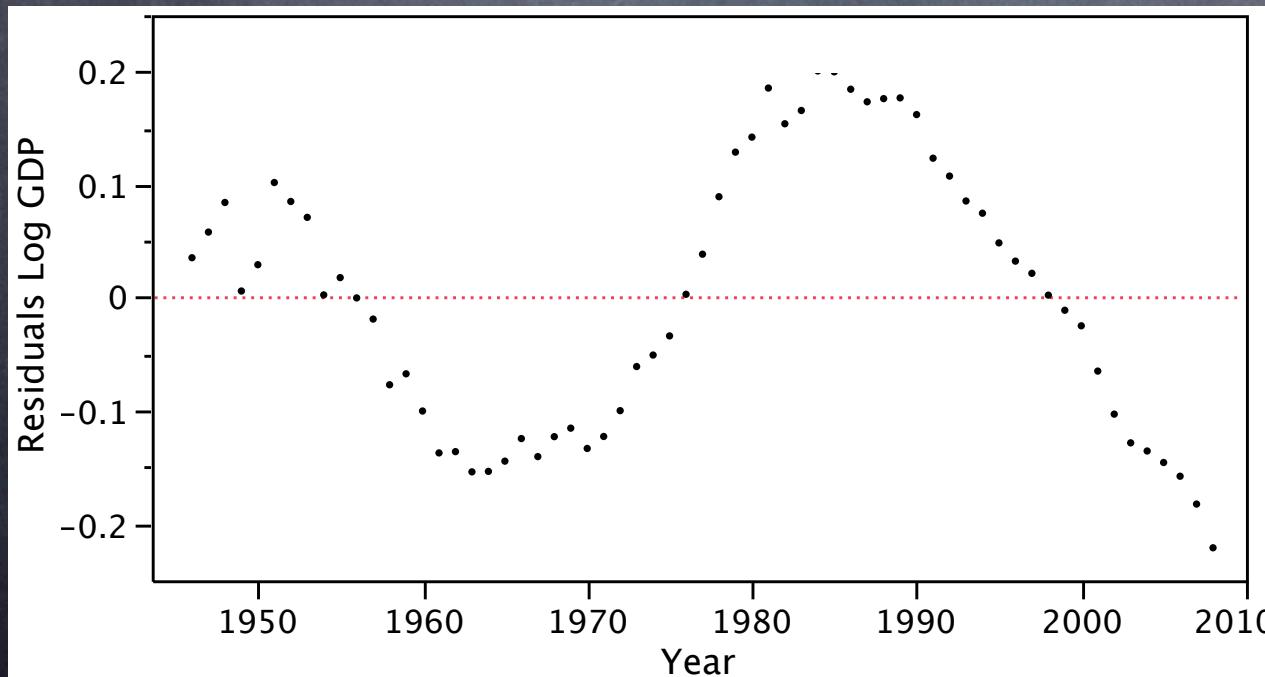


Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	-133.3041	1.629637	-81.80	<.0001*
Year	0.0712604	0.000824	86.45	<.0001*

Concerns?

# GDP Residuals

- Pattern in residuals from linear regression
  - Residuals suggest stationary time series
  - Meandering pattern implies dependence
  - Oscillates around fit of model on log scale
    - Regression over-predicts GDP since 2000
  - Autocorrelation between adjacent residuals  
 $\text{corr}(e_t, e_{t-1}) \approx 0.94$



Durbin-Watson		
Durbin-Watson	Number of Obs.	AutoCorrelation
0.0596695	63	0.9411

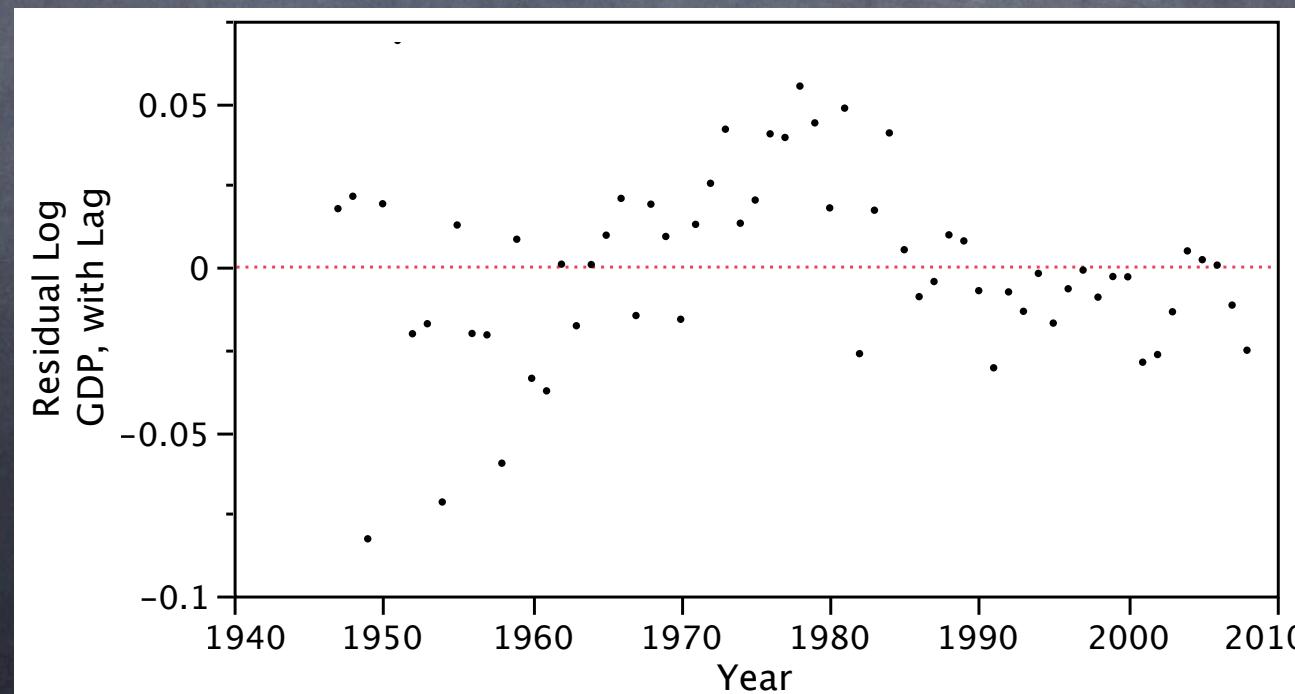
Look back:  
Pattern becomes evident  
in scatterplot of GDP on  
time once you have seen  
it in the residual plot.

# Improved Model

- Follow Durbin-Watson approach
  - Add lag of residuals back into equation
  - Refit regression with lagged variable
- Multiple regression with lagged variable
  - Very high  $R^2$  (0.9995)
  - But DW remains significant
  - Does more dependence remain in the residuals?

Term	Estimate	Std Error	t Ratio
Intercept	-132.7244	0.403924	-328.6
Year	0.0709652	0.000204	347.43
Lag Log Residuals GDP	1.0013018	0.031908	31.38

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
1.5222692	62	0.2290	0.0143*



# Closer Look At DW Stat

- Durbin-Watson statistic tests for a specific type of dependence among residuals:  
first-order autocorrelation
- Model: First-order autoregression AR(1)
  - Equation
$$(y_t - \mu) = \varphi(y_{t-1} - \mu) + a_t \quad \mu = E(y_t)$$
  - Errors  $a_t$  have mean zero, equal variance, independent
    - $E(a_t) = 0, \quad \text{Var}(a_t) = \sigma^2, \quad \text{Corr}(a_t, a_s) = 0, \quad s \neq t$
- Questions
  - How do we identify whether the dependence among the errors in our model is AR(1) or some other type?
  - What other types are there?

# Recognizing AR(1) Dependence

- A first-order autoregression is a geometrically weighted sum of past error terms.

- Assume  $|\varphi| < 1$
- Back-substitute, assuming  $\mu = 0$  (e.g. residuals)

$$\begin{aligned}y_t &= \varphi y_{t-1} + a_t \\&= \varphi(\varphi y_{t-2} + a_{t-1}) + a_t \\&= \varphi(\varphi(\varphi y_{t-3} + a_{t-2}) + a_{t-1}) + a_t \\&= a_t + \varphi a_{t-1} + \varphi^2 a_{t-2} + \varphi^3 y_{t-3} \\&= a_t + \varphi a_{t-1} + \varphi^2 a_{t-2} + \dots + \varphi^m a_{t-m} + \dots\end{aligned}$$

- Autocorrelation decays

- First two

$$\text{Corr}(y_t, y_{t-1}) = \text{Corr}(\varphi y_{t-1} + a_t, y_{t-1}) = \varphi$$

$$\text{Corr}(y_t, y_{t-2}) = \text{Corr}(\varphi^2 y_{t-2} + \varphi a_{t-1} + a_t, y_{t-1}) = \varphi^2$$

- In general

$$\text{Corr}(y_t, y_{t-k}) = \varphi^k, k = 1, 2, 3, \dots$$

# Autocorrelation Function

## • Identification procedure

- Autocorrelations decay geometrically if process is AR(1)
- Plot the autocorrelations; decide if decay looks geometric.
- Notation

$$TAC(k) = \rho_k = \text{Correlation}(Y_t, Y_{t-k}) \quad [\text{a.k.a., ACF}]$$

## • Complications

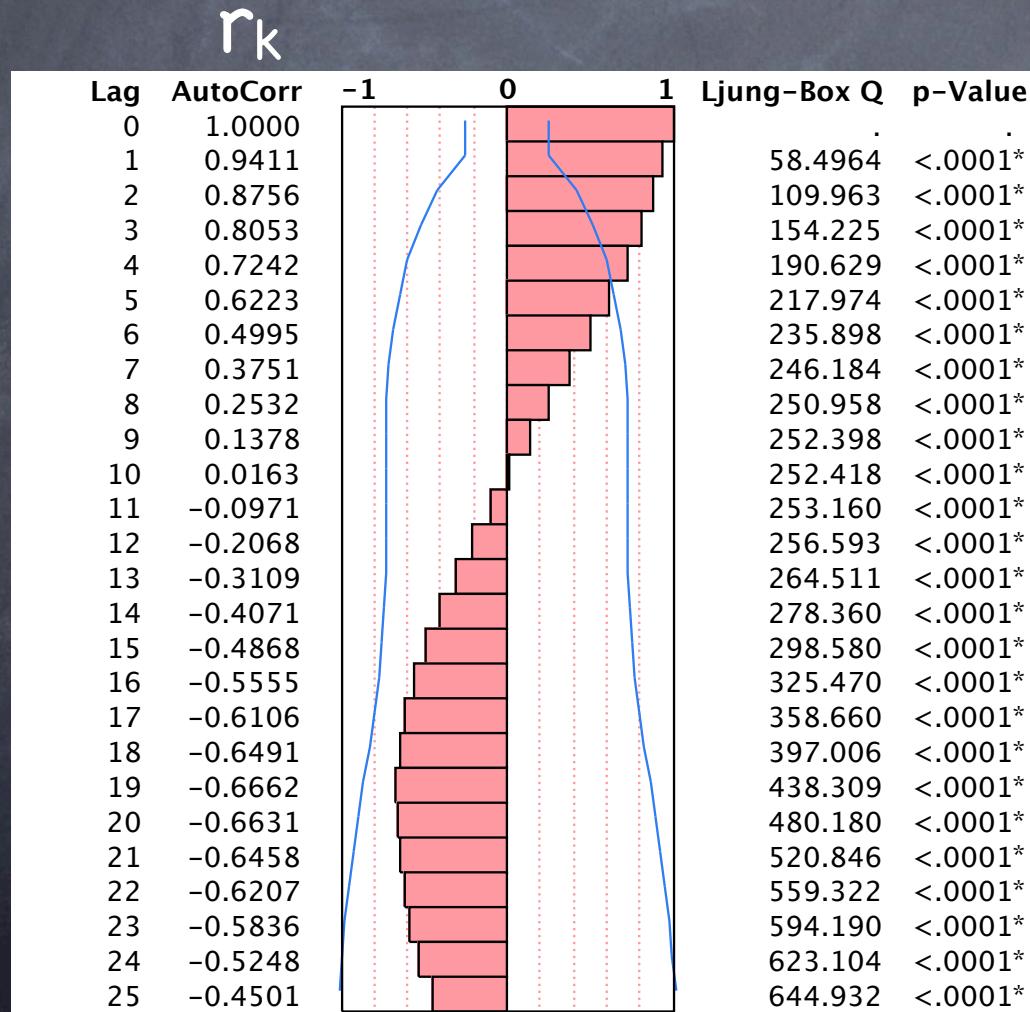
- We observe sample correlations (**SAC**  $r_k$ ), not true values
- Outliers, nonlinearity can make a mess of things
- How close to geometric is close enough?

## • Calculation

- JMP plots the autocorrelation function for you without having to form all of the lags (beware outliers!)
- Analyze > Modeling > Time Series

# GDP Residual ACF

- SAC starts out positive, then drifts negative
  - Not a geometric decay toward 0
  - Values start out positive, then gradually become negative



$H_0:$   
No cumulative  
autocorrelation

Conclude: Not  
AR(1) process.

# Models of Dependence

- ⦿ “Box-Jenkins” family of models (ARIMA models)
  - ⦿ Named for authors of influential book
- ⦿ Autoregressions      AR( $p$ ),  $p = 1, 2, \dots$ 
  - ⦿ Regression models that describe the current value of the time series  $Y_t$  as a weighted sum of  $p$  past values of the time series.
- ⦿ Moving averages      MA( $q$ ),  $q = 1, 2, \dots$ 
  - ⦿ Represent the time series as a weighted sum of  $q$  unobserved, uncorrelated error terms.
- ⦿ Autoregressive, moving-average (ARMA) combine the two types of models

# Distinguishing AR(p) Models

- ⦿ AR(p) model adds lags of the time series
  - $Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + a_t \quad (\mu=0)$
  - Stationarity constrains the coefficients
    - Analogous to keeping  $|\varphi| < 1$  in AR(1) model
- ⦿ Complication:  
All AR(p) models have geometric decay in TAC
  - How do we distinguish an AR(2) from an AR(4)?
  - You cannot, at least not based on estimates.
- ⦿ Distinguishing among AR(p) models requires another plot
  - Need to determine reasonable choice for p
  - Partial autocorrelation function (TPAC or SPAC, PACF)

# Partial Autocorrelation

## • Partial correlation

- Correlation between two random variables conditional on a collection of other variables

$$\text{Corr}(X, Y | Z_1, Z_2, \dots, Z_k)$$

- Example: multiple regression of  $Y$  on  $X$  and  $Z_1, Z_2, \dots, Z_k$ .
- Leverage plot is a plot that shows the partial correlation between  $Y$  and  $X$  having adjusted for  $Z_1, Z_2, \dots, Z_k$ .

## • Role in autoregression

- Suppose  $Y_t$  is an AR(2) process

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + a_t$$

- Consider partial correlation conditional on  $Y_{t-1}$  and  $Y_{t-2}$ .

$$\text{Corr}(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2})$$

- What's this going to be if  $k > 2$ ?

- Construction of leverage plots provides a hint.

# Graphing PACF

- Construct the sequence

$$\text{Corr}(Y_t, Y_{t-1})$$

$$Y_t = \varphi_1 Y_{t-1} + a_t$$

$$\text{Corr}(Y_t, Y_{t-2} | Y_{t-1})$$

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + a_t$$

$$\text{Corr}(Y_t, Y_{t-3} | Y_{t-1}, Y_{t-2})$$

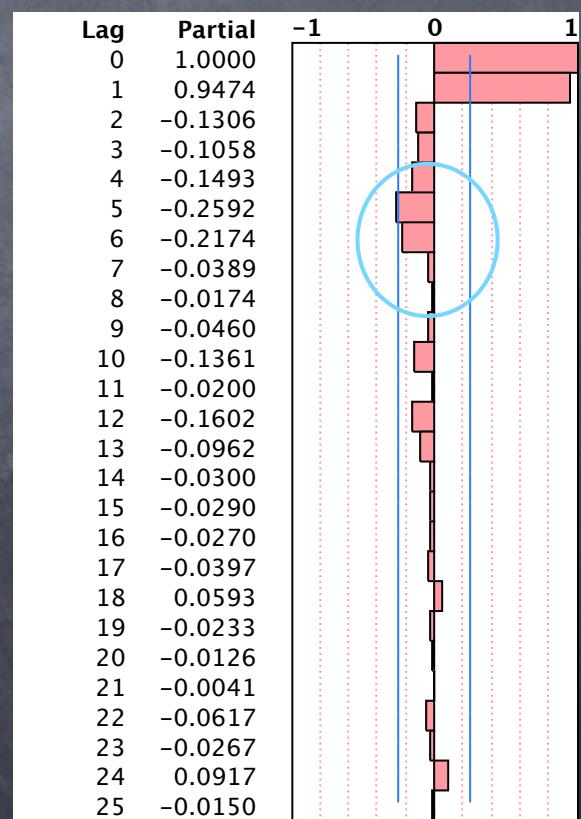
$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + a_t$$

...

$$\text{Corr}(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \dots, Y_{t-k-1})$$

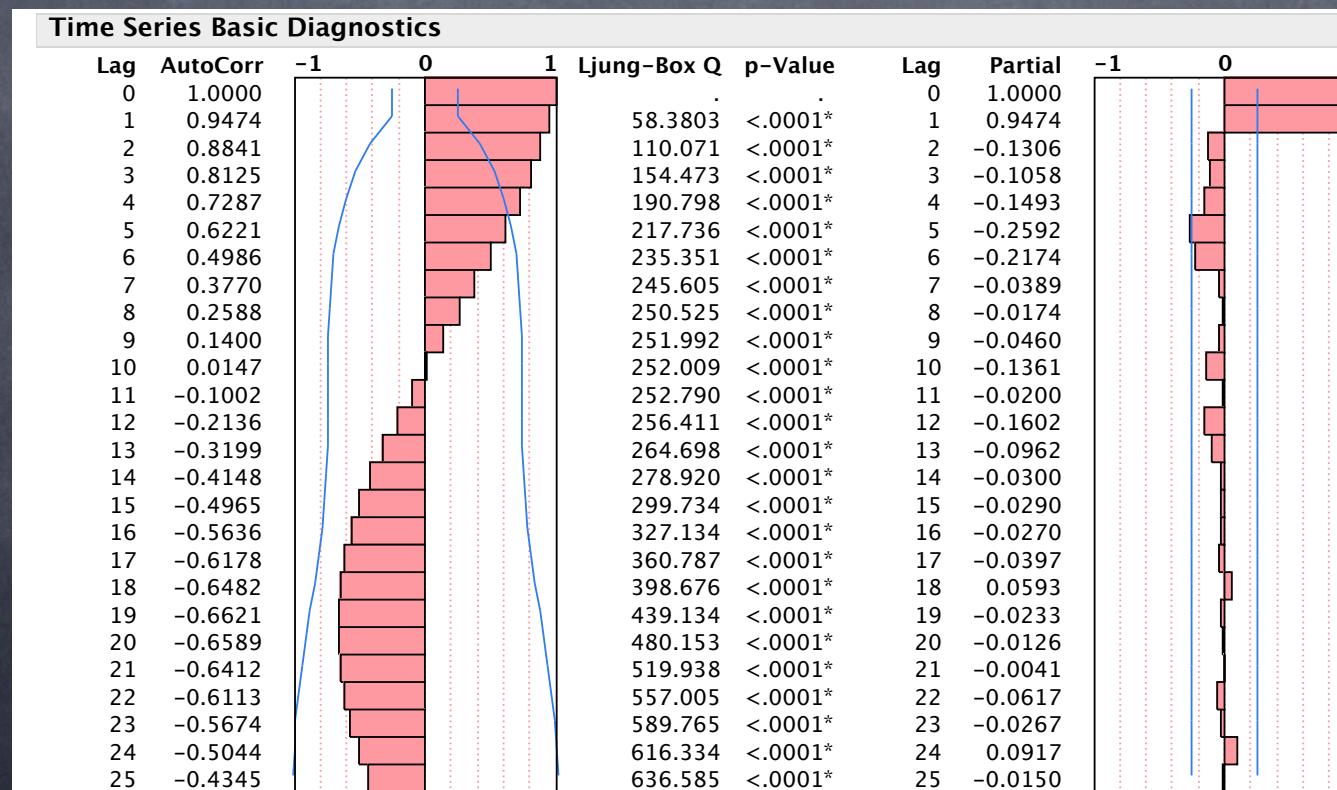
- Graph the sequence of estimates versus lag k

- Denote estimates  $r_{kk}$
- Dependence at separation of about 5-6 years indicates need for a more complex AR model



# JMP Software

- Time series platform shows SAC, SPAC side-by-side
- Autoregression
  - Geometric decay in the SAC
  - Sharp “cut-off” in the SPAC



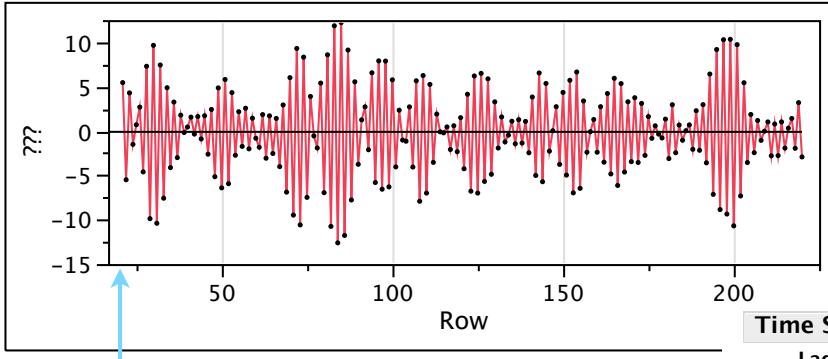
$r_k$

$r_{kk}$

Fact that we observe estimates rather than population values complicates task.

# Simulated Example

- Generate data from an AR process... Which is it?



Exclude the first 20 rows for "burn in", start the recursion

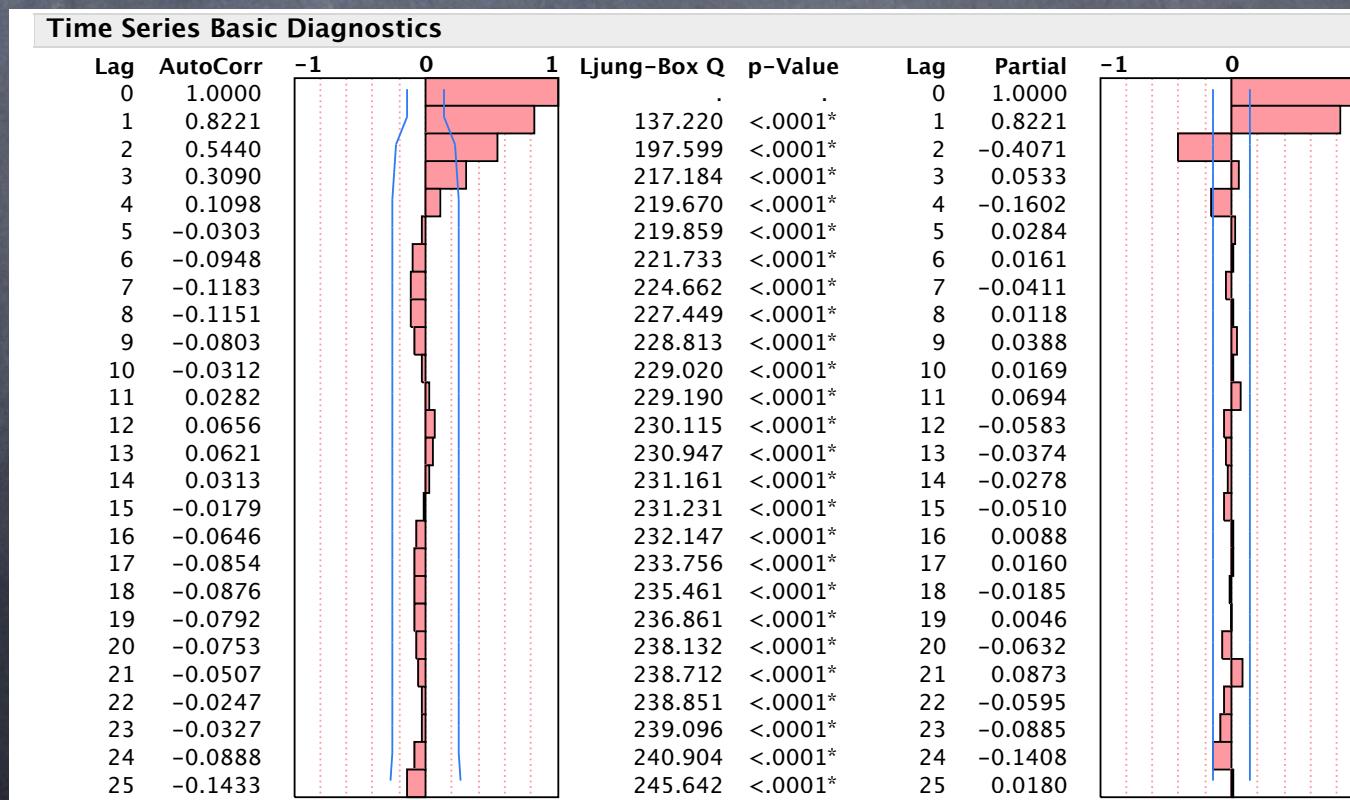
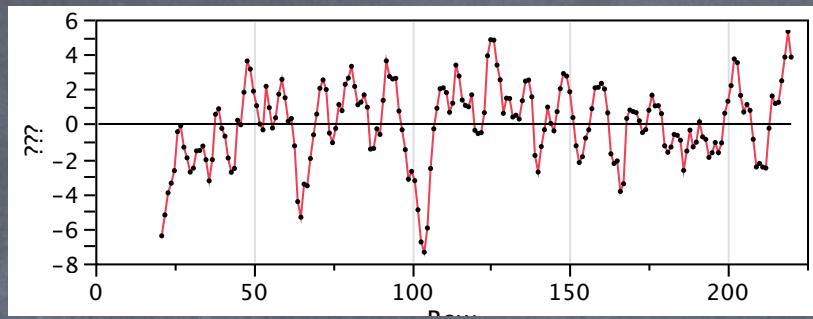
What happens if observe fewer data than these 200?

Time Series Basic Diagnostics

Lag	AutoCorr	-1	0	1	Ljung-Box Q	p-Value	Lag	Partial	-1	0	1
0	1.0000				.	.	0	1.0000			
1	-0.9499				183.180	<.0001*	1	-0.9499			
2	0.8730				338.677	<.0001*	2	-0.3001			
3	-0.7362				449.830	<.0001*	3	0.6391			
4	0.5888				521.291	<.0001*	4	0.0165			
5	-0.4187				557.603	<.0001*	5	-0.0249			
6	0.2502				570.636	<.0001*	6	-0.1301			
7	-0.0836				572.098	<.0001*	7	0.0389			
8	-0.0745				573.268	<.0001*	8	-0.1296			
9	0.2110				582.684	<.0001*	9	-0.0560			
10	-0.3319				606.105	<.0001*	10	-0.0640			
11	0.4200				643.810	<.0001*	11	-0.0033			
12	-0.4894				695.274	<.0001*	12	-0.0666			
13	0.5257				754.976	<.0001*	13	0.0334			
14	-0.5371				817.639	<.0001*	14	0.0956			
15	0.5234				877.471	<.0001*	15	0.0878			
16	-0.4833				928.748	<.0001*	16	0.0508			
17	0.4249				968.596	<.0001*	17	-0.0414			
18	-0.3485				995.556	<.0001*	18	-0.0660			
19	0.2609				1010.75	<.0001*	19	-0.0801			
20	-0.1637				1016.76	<.0001*	20	0.0274			
21	0.0615				1017.61	<.0001*	21	-0.0610			
22	0.0407				1017.99	<.0001*	22	0.0188			
23	-0.1434				1022.68	<.0001*	23	-0.0859			
24	0.2276				1034.57	<.0001*	24	-0.1442			
25	-0.3106				1056.84	<.0001*	25	-0.1572			

# Another Example

• What about this time series?



# Moving Average Models

## • Moving Average models MA(q)

- Weighted average of error terms

- MA(q)  $Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$

## • Dual behavior compared to AR models

- TAC  $\rho_k$

Cuts off since errors are uncorrelated once the observations are more than q terms apart.

- TPAC  $\rho_{kk}$

Geometric decay

- Estimates may not look much like either

## • Relationship to AR models

- AR models are an infinite weighted average of past  $a_t$

- MA models only depend on a finite number of  $a_t$ .

# ARMA Models

- ⦿ Mixture of autoregressive and moving average components

- ⦿ ARMA(p,q)

$$(Y_t - \mu) = \varphi_1 (Y_{t-1} - \mu) + \dots + \varphi_p (Y_{t-p} - \mu) \\ + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

- ⦿ Correlation functions

- ⦿ TAC

Geometric decay

- ⦿ TPAC

Geometric decay

- ⦿ Identification procedure is iterative

- ⦿ Guess from substantive knowledge, SAC, and SPAC

- ⦿ Estimate the model to see how well it does

# Summary

- AR, MA and ARMA models for stationary series
  - Particularly useful models for residual processes

- Correlation functions

Fig 9.5-9.6, page 412

	AR(p)	MA(q)	ARMA
TAC	geometric	cuts off	geometric
TPAC	cuts off	geometric	geometric

- Challenges that remain
  - Estimation
  - Diagnostics
  - Prediction

What to use to predict GDP?