APS 425 – Winter 2008

Time Series Analysis: Trend & Seasonality

Instructor: G. William Schwert
585-275-2470
schwert@schwert.simon.rochester.edu

Topics

- Deterministic trends & seasonality
- Stochastic trends & seasonality
  - Differencing & seasonal differencing

Time Trends

Linear

$Z_t = a + b\ t + e_t$

- do you really expect such a pattern to persist forever?
- e.g., accumulated horse manure in the streets of Philadelphia in the late 18th century

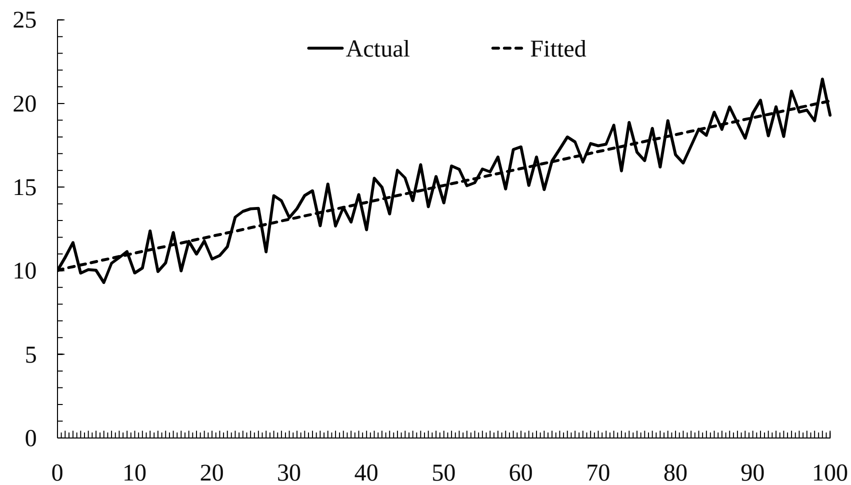
exponential (linear in logs)

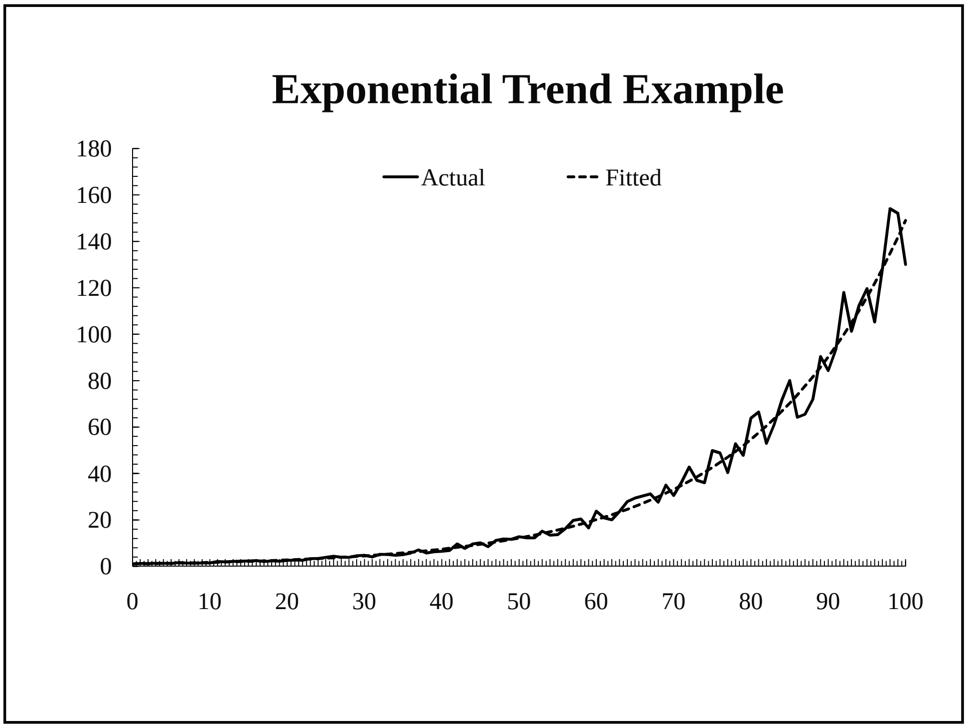
$\log(Z_t) = a + b\ t + e_t$

$Z_t = \exp\{a + b\ t + e_t\}$

In Eviews, time = @trend(K) creates a time trend equal to 0 in observation K

Linear Trend Example



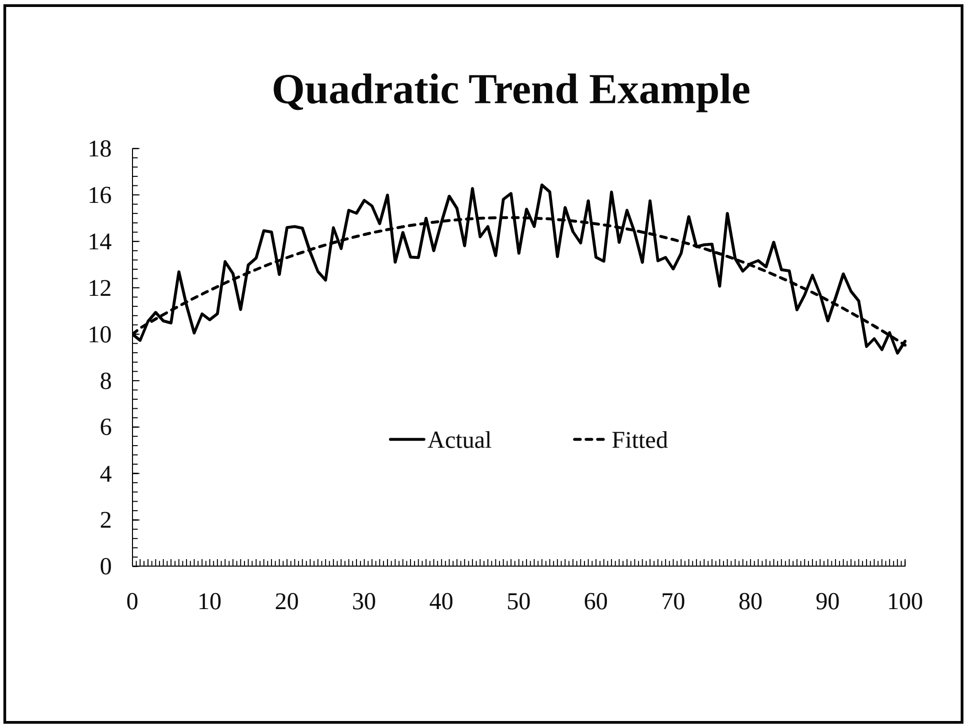


Polynomial Trends

$Z_t = a + b_1 t + b_2 t^2 + e_t$

- parabola
  - if $b_2 < 0$, then it eventually turns negative
  - if $b_2 > 0$, then it eventually sky-rockets up

- you can approximate most samples with polynomials, but you wouldn't expect to do well forecasting out-of-sample



Other Functions of Time

Cosine or Sine represent fixed cycles

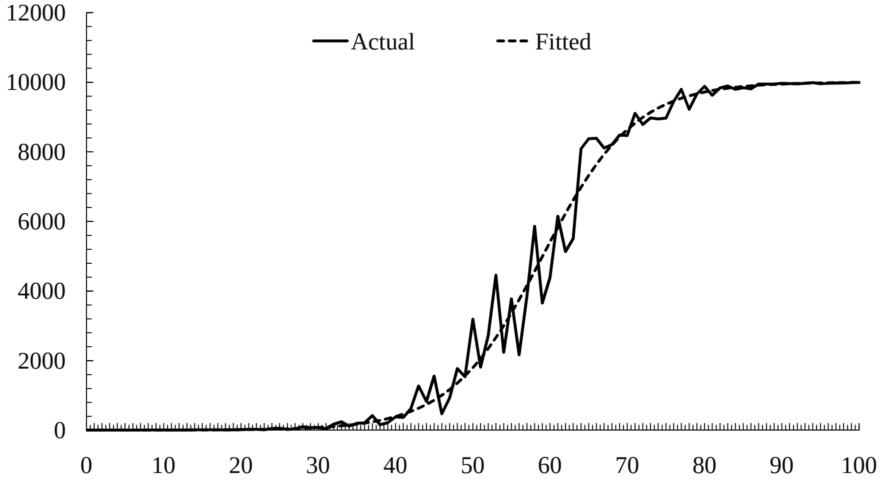
S-curves (logistic) -- new product sales

$\log((L - Z_t)/Z_t) = a + b\,t + e_t$

$Z_t = L / \{1 + \exp[\,a + b\,t + e_t]\}$

- L is the upper limit (asymptote) for sales
- can iterate on this by trying different values to see which fits best

S-Curve Trend Example



Other Types of Stationarity-inducing Transformations

- per capita
  - divide by population (if it is growing)
- real
  - divide by price index (e.g., CPI)
- market share
  - divide by industry sales

Relation to Differencing

Linear Trend:

$Z_t = a + b\,t + e_t$

$(Z_t - Z_{t-1}) = [a + b\,t + e_t] - [a + b\,(t-1) + e_{t-1}]$

$= b + e_t - e_{t-1}$

Random Walk with Drift:

$Z_t = b + e_t + Z_{t-1}$

$(Z_t - Z_{t-1}) = b + e_t$

The Level of a Random Walk with Drift

$Z_t = b + e_t + Z_{t-1}$

$= b + e_t + b + e_{t-1} + Z_{t-2}$

$= b + e_t + b + e_{t-1} + b + e_{t-2} + Z_{t-3}$

. . .

$= b\,t + e_t + e_{t-1} + e_{t-2} + \ldots + e_1 + Z_0$

- which is the initial value of Z_0,
  plus b times t,
  plus the sum of all the shocks from time 1 through t

- so Z_0 plays the role of “a” in the trend model, but the errors follow a random walk

Seasonality

Any type of predictable behavior that is a function of the "season"

- originally used for agricultural products
- growing/harvests happen at the same time every year

Also caused by conventions of holidays, etc.

- e.g., more sales at Christmas

Can also reflect days of the week

- e.g., Fed Funds rate

Seasonality

1) Fixed effects -- dummy variables

2) ARMA model at seasonal frequencies

Seasonals (Fixed)

Monthly dummy variables
- e.g., if Jan = 1; otherwise = 0
- allows for different intercepts by month

Day-of-the-week dummy variables for daily data
- also, for holidays

In Eviews
- Jan = @seas(1) if the workfile is defined as being monthly

Seasonals (Fixed) Monthly Example

define Jan_t = 1 if the observation occurs in January, and = 0, otherwise; and so forth

$$Z_t = \alpha_1 Jan_t + \alpha_2 Feb_t + \ldots + \alpha_{12} Dec_t + e_t$$

so α_1 is the average value of Z in Januarys, etc.

This is equivalent to the model where any one of the monthly dummies is left out (e.g., Jan)

$$Z_t = \alpha + \alpha_2 Feb_t + \ldots + \alpha_{12} Dec_t + e_t$$

so α_2 is the average difference between the mean in February and the mean in January (which is α)

Seasonal Differencing

Similar to ordinary differences, except s periods apart (Z_t - Z_{t-s}):

- note that seasonal differencing removes a linear trend
- it also removes different fixed means
  - i.e., dummy variables

Relations Among Seasonal Variables

Suppose you have a regression model that relates sales to past advertising for a product that has a seasonal sales cycle -- e.g., automobiles

- it is easy to imagine that some types of advertising that are specific to a particular model or new feature would have no effect on sales beyond this model year
- other types of advertising that highlighted general quality characteristics of the brand might have effects that last beyond the model year
- when building a model you need to <u>think</u> about what effects seasonality might have on the model

Seasonal ARIMA$(P,D,Q)_s$ Models

Just like ARIMA(p,d,q) models, except the gap between observations that affect one another is s periods, not 1 period

Example: ARIMA$(1,0,0)_{12}$ model [monthly AR(1)]

$$Z_t = \alpha + \phi_1 Z_{t-12} + e_t$$

- so the autocorrelations decay exponentially at lags 12, 24, 36, etc.
- partial autocorrelation at lag 12 = ϕ_1
  - after lag 12, they equal 0

Seasonal ARIMA$(0,1,1)_s$ Models

This model occurs a lot in real data

- note that seasonal differencing removes a linear trend
- it also removes different fixed means
  - i.e., dummy variables
- if the MA parameter is close to 1, this "exponential smoothing" model implies a slowly wandering level for the series that is different for each period of the year

Seasonal Exponential Smoothing Forecasts

ARIMA$(0,1,1)_{12}$ model:

$$[Z_t - Z_{t-12}] = e_t - \theta_1 e_{t-12}$$

$$\hat{Z}_{t-12}(12) = Z_{t-12} - \theta_1 e_{t-12} = Z_{t-12} - \theta_1 [Z_{t-12} - \hat{Z}_{t-24}(12)]$$

$$= (1 - \theta_1) Z_{t-12} + \theta_1 \hat{Z}_{t-24}(12)$$

- so the forecast for next January is a weighted average of the most recent observation and the forecast for that observation (for Januarys)

Seasonal ARIMA$(P,D,Q)_S$ Models Interacting with ARIMA(p,d,q) Models

This is the "multiplicative" seasonal ARIMA model:

$$\Phi_P(L^S)\ \phi_p(L)\ [1-L^S]^D [1-L]^d\ Z_t = \alpha + \Theta_Q(L^S)\ \theta_q(L)\ e_t$$

- if you want to figure out what pattern of autocorrelations are implied, multiply out the two autoregressive polynomials and the two moving average polynomials, and also the difference operators

Seasonal ARIMA(P,D,Q)$_S$ Models Interacting with ARIMA(p,d,q) Models

Example -- ARIMA(0,1,1)$_{12}$ ARIMA(0,0,1):

$$Z_t = \alpha + (1 - \Theta L^{12})(1 - \theta L)\, e_t$$

$$= \alpha + (1 - \theta L - \Theta L^{12} + \theta \Theta L^{13})\, e_t$$

- which is equivalent to an ARIMA(0,0,13) model with the following parameters:
  - $\theta_1 = \theta$
  - $\theta_k = 0$, for $1 < k < 12$
  - $\theta_{12} = \theta$; $\theta_{13} = -\theta\Theta$
  - $\theta_k = 0$, for $k > 13$

Seasonal Adjustment

The most common method is the X-11 technique developed at the U.S. Bureau of the Census

- a complex (two-sided) filter that involves ratios and moving averages
- I have seen cases (e.g., CPI) where it does not seem to work well
- for forecasting purposes, it is difficult, since the final seasonally adjusted figure will keep getting revised as new data comes in
  - it is a two-sided filter

Seasonality: Summary

1) If the most interesting variation in the data is at seasonal lags, then you probably need more data to get a reliable model
   - e.g., with monthly dummies, you are trying to estimate the monthly averages with only T/12 observations per parameter

2) It is probably better to directly model seasonality, rather than use seasonally adjusted data
   - mechanical adjustment procedures often don't work well

Links

Return to APS 425 Home Page:

http://schwert.simon.rochester.edu/a425/a425main.htm