Time Series Forecasting Methods

Nate Derby

Statis Pro Data Analytics
Seattle, WA, USA

Calgary SAS Users Group, 11/12/09

Outline

1. Introduction
   - Objectives
   - Strategies
2. Univariate Forecasting
   - Seasonal Moving Average
   - Exponential Smoothing
   - ARIMA
3. Conclusions
   - Which Method?
   - Are Our Results Better?
   - What's Next?

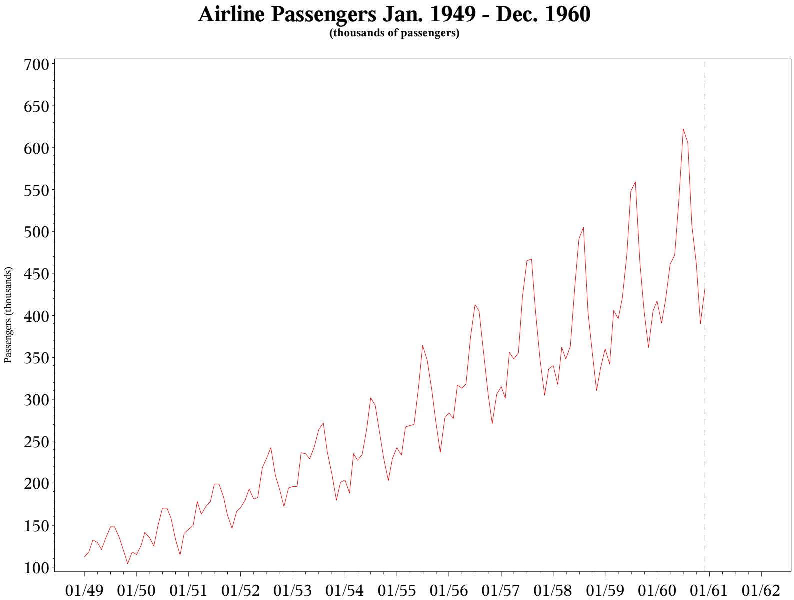
Objectives

- What is time series data?
- What do we want out of a forecast?
  - Long-term or short-term?
  - Broken down into different categories/time units?
  - Do we want *prediction intervals*?
  - Do we want to measure effect of X on Y? (scenario forecasting)
- What methods are out there to forecast/analyze them?
- How do we decide which method is best?
- How can we use SAS for all this?

What is Time Series Data?

Time Series data = Data with a pattern ("trend") over time.

- Ignore time trend = Get wrong results.
    - See my `PROC REG` paper.


Airline Passengers Jan. 1949 - Dec. 1960
(thousands of passengers)
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62

Base Data Set

	date	pass
1	01/49	112
2	02/49	118
3	03/49	132
4	04/49	129
5	05/49	121
6	06/49	135
7	07/49	148
8	08/49	148
9	09/49	136
10	10/49	119
11	11/49	104
12	12/49	118
13	01/50	115
14	02/50	126
15	03/50	141

What do we want out of a Forecast?

- Long-term:
  - Involves *many* assumptions! (e.g., global warming)
  - Involves tons of uncertainty.
  - Keynes: “In the long run we are all dead”.
  - We’ll focus on the short term.
- Different categories?
  - Two strategies for forecasting A, B and C:
    1. Forecast their combined total, then break it down by percentages.
    2. Forecast them separately.
  - Idea: Do (1) unless percentages are unstable.

What do we want out of a Forecast?

- Different time units?
  - Two strategies for forecasting at two different time units (e.g., daily and weekly):
    1. Forecast weekly, then break down into days by percentages.
    2. Forecast daily, then aggregate into weeks.
  - Idea: Idea: Do (1) unless percentages are unstable.
- Do we want prediction intervals?
  - Prediction interval = Interval where data point will be with 90/95/99% probability.
  - Yes, we want them!

What do we want out of a Forecast?

- Do we want to measure effect of X on Y?
  - Ex: Marketing campaign $\Rightarrow$ calls to call center.
  - Harder to do, but
  - Allows for scenario forecasting!
  - Idea: Do it, but only with most important Xs.

Remaining Questions: Basis of this talk:

- What methods are out there to forecast/analyze them?
- How do we decide which method is best?
- How can we use SAS for all this?
  - Methods will require `ETS` package.

Strategies

Two stages:

- *Univariate* (one variable) forecasting:
  - Forecasts Y from trend alone.
  - Gives us a basic setup.
- *Multivariate* (many variables) forecasting:
  - Forecasts Y from trend and other variables $X_1, X_2, \ldots$.
  - Allows for "what if" scenario forecasting.
  - May or may not make more accurate forecasts.

Univariate Forecasting - Intro

- Gives us a benchmark for comparing multivariate methods.
- Could give better forecasts than multivariate.
- Some methods can be extended to multivariate.
- Currently three methods:
  - Seasonal moving average (*very* simple)
  - Exponential smoothing (simple)
  - ARIMA (complex)
- More complex methods, for later on (for me):
  - State space (promising)
  - Bayesian (maybe . . .)
  - Wavelets? (forget it!)

Once Again ...

Q: Why not use `PROC REG`?

$$Y_t = \beta_0 + \beta_1 X_t + Z_t$$

- A: We can get misleading results (see my `PROC REG` paper).

Seasonal Moving Average

Simple but sometimes effective!

- Moving Average:

$$\text{Forecast} = \text{Average of last } n \text{ months.}$$

- Seasonal Moving Average:

$$\text{Forecast} = \text{Average of last } n \text{ Novembers.}$$

- After a certain point, forecast the same for each of same weekday.
  - Doesn't allow for a trend.
- Not based on a *model* $\Rightarrow$ No prediction intervals.

SAS Code

Making lags in a `DATA` step (to make the averages) is not fun:

Making 4 lags (Brocklebank and Dickey, p. 45)

```
DATA movingaverage;
  ...
  RETAIN date pass1-pass4;
  OUTPUT;
  pass4=pass3;
  pass3=pass2;
  pass2=pass1;
  pass1=pass;
RUN;
```

SAS Code

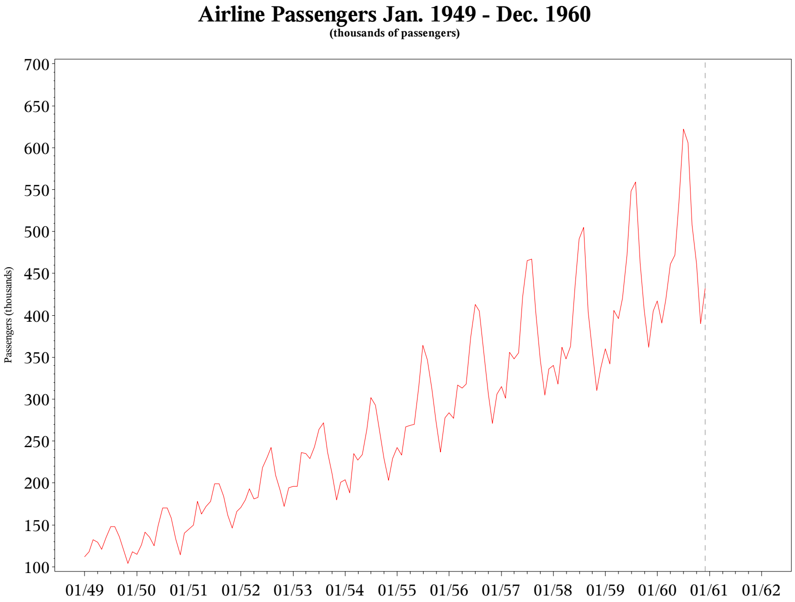
Much easier with a trick with `PROC ARIMA`.

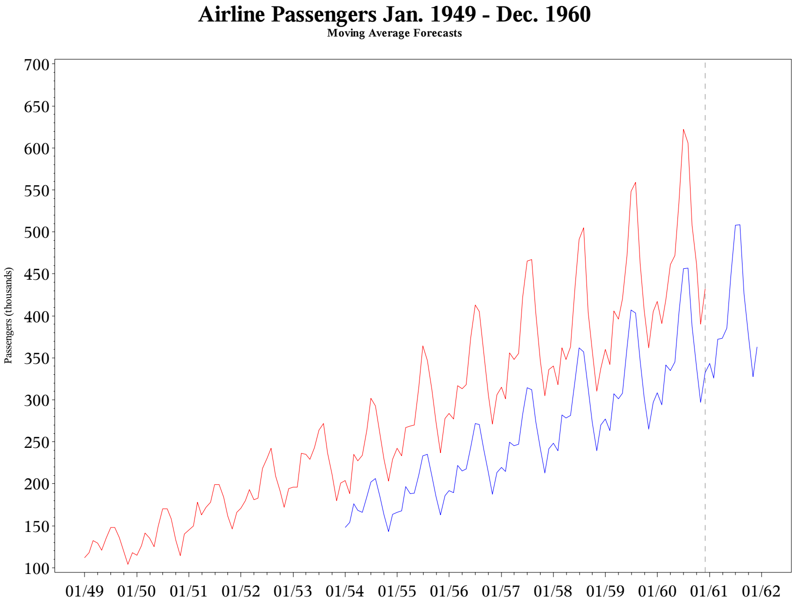
Seasonal = averaging over past 5 years on that same month:

$$Y_t = \frac{1}{5}\left(Y_{t-12} + Y_{t-24} + Y_{t-36} + Y_{t-48} + Y_{t-60}\right)$$

Forecasting 3 weeks ahead, seasonal moving average

```
PROC ARIMA data=airline;
  IDENTIFY var=pass noprint;
  ESTIMATE p=( 12, 24, 36, 48, 60 ) q=0 ar=0.2 0.2 0.2 0.2 0.2
    noest noconstant noprint;
  FORECAST lead=12 out=foremave id=date interval=month noprint;
RUN;
QUIT;
```


Airline Passengers Jan. 1949 - Dec. 1960
(thousands of passengers)
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62


Airline Passengers Jan. 1949 - Dec. 1960
Moving Average Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62

Exponential Smoothing I

Notation: $\hat{y}_t(h)$ = forecast of Y at horizon h, given at time t.

- Idea 1: Predict Y_{t+h} by taking weighted sum of past observations:

$$\hat{y}_t(h) = \lambda_0 y_t + \lambda_1 y_{t-1} + \cdots$$

  Assumes $\hat{y}_t(h)$ is constant for all horizons h.

- Idea 2: Weight recent observations heavier than older ones:

$$\lambda_i = c\alpha^i,\ 0 < \alpha < 1 \ \Rightarrow\ \hat{y}_t(h) = c\left(y_t + \alpha y_{t-1} + \alpha^2 y_{t-2} + \cdots\right)$$

  where c is a constant so that weights sum to 1.

Exponential Smoothing II

$$\hat{y}_t(h) = c\left(y_t + \alpha y_{t-1} + \alpha^2 y_{t-2} + \cdots\right)$$

- Weights are *exponentially decaying* (hence the name).
- Choose α by minimizing squared one-step prediction error.

Overall:

- Just a weighted moving average.
- Can be extended to include trend and seasonality.
- Prediction intervals? Sort of ...

SAS Code

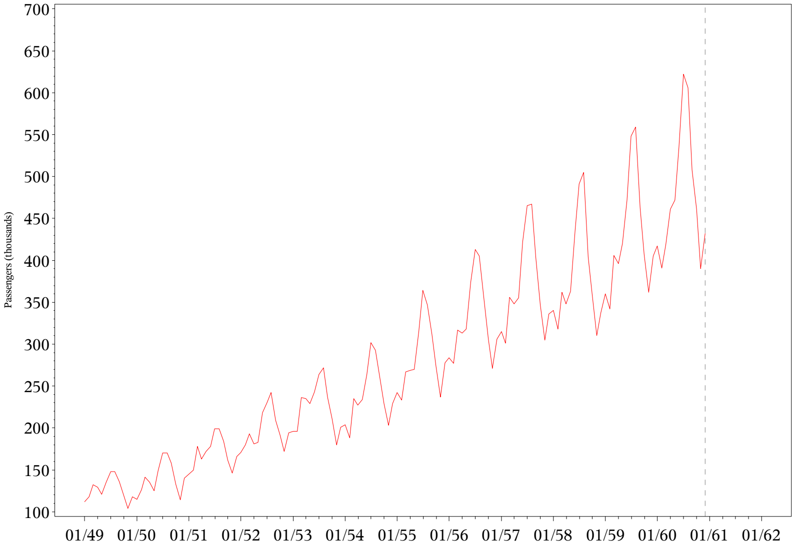
All done with `PROC FORECAST`:

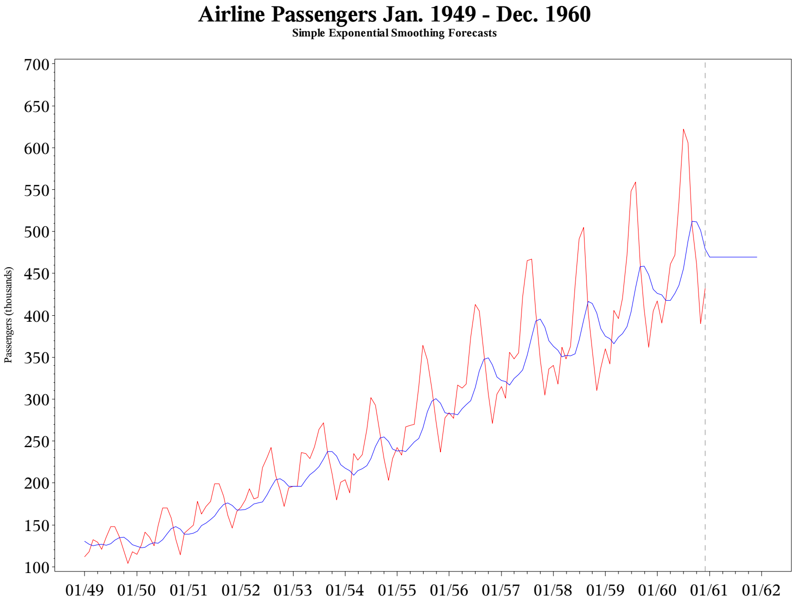
- `method=expo trend=1` for simple.
- `method=expo trend=2` for trend.
- `method=winters seasons=( 12 )` for seasonal.

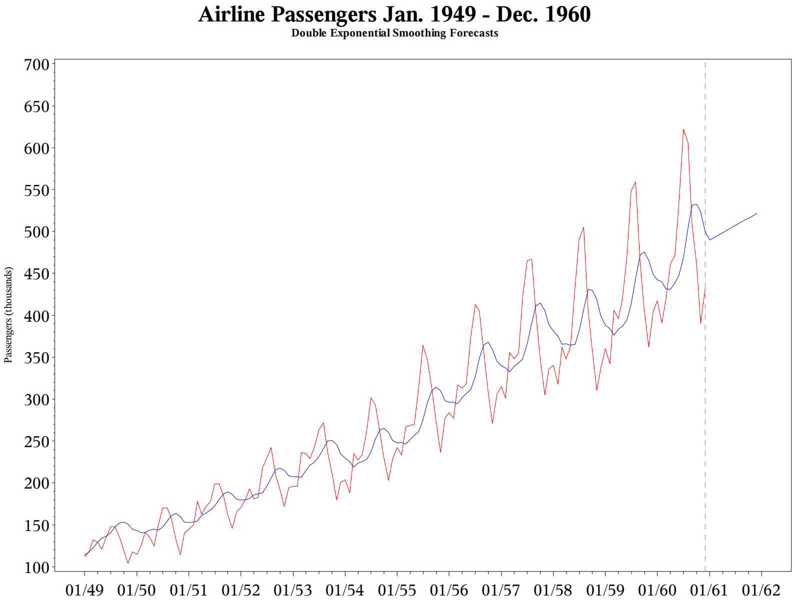
Forecasting 3 weeks ahead, exponential smoothing

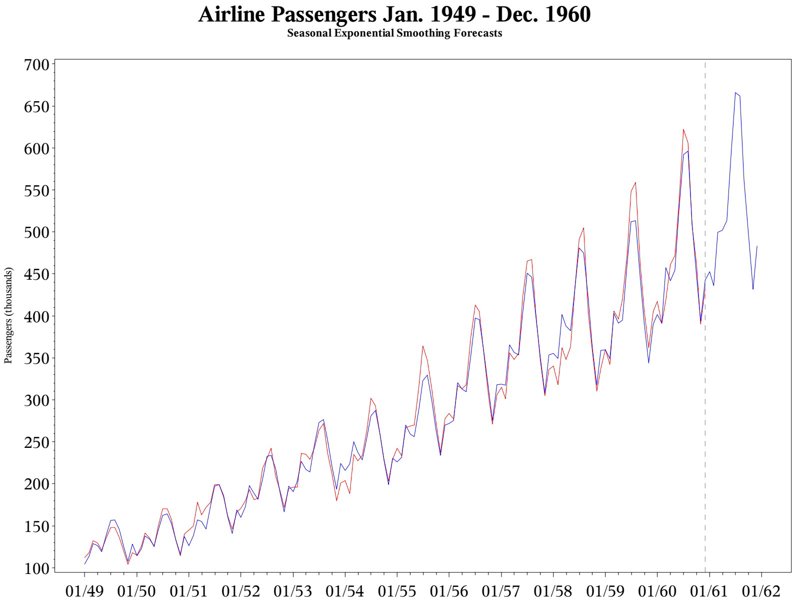
```
PROC FORECAST data=airline method=xx interval=month lead=12
    out=foreexsm outactual out1step;
  VAR pass;
  ID date;
RUN;
```

Airline Passengers Jan. 1949 - Dec. 1960
(thousands of passengers)




Airline Passengers Jan. 1949 - Dec. 1960
Simple Exponential Smoothing Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62


Airline Passengers Jan. 1949 - Dec. 1960
Double Exponential Smoothing Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62


Airline Passengers Jan. 1949 - Dec. 1960
Seasonal Exponential Smoothing Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62

Exponential Smoothing VI

Advantages:

- Gives interpretable results (trend + seasonality).
- Gives more weight to recent observations.

Disadvantages:

- Not a model (in the statistical sense).
  - Prediction intervals not (really) possible.
- Can't generalize to multivariate approach.

ARIMA I

- Stands for *AutoRegressive Integrated Moving Average* models.
- Also known as Box-Jenkins models (Box and Jenkins, 1970).
- Advantages:
  - Best fit (minimum mean squared forecast error).
  - Generalizes to multivariate approach.
  - Often used in *statistical* practice.
- Disadvantages:
  - More complex.
  - Not intuitive *at all*.

ARIMA II

Assume nonseasonality for now.

- First, transform, then difference the data $\{Y_t\}$ d times until it is stationary (constant mean, variance), denoted $\{Y_t^*\}$.
- Guesstimate orders p, q through the sample autocorrelation, partial autocorrelation functions.
- Fit an *autoregressive moving average* (ARMA) process, orders p and q:

$$\begin{aligned} Y_t^* - \phi_1 Y_{t-1}^* - \cdots - \phi_p Y_{t-p}^* &= Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \\ \phi(Y_t^*) &= \theta(Z_t) \end{aligned}$$

  where $Z_t \overset{iid}{\sim} N(0, \sigma^2)$, and $\phi_1, \ldots, \phi_p$, $\theta_1, \ldots, \theta_q$ are constants.
- Through trial and error, repeat above 2 steps until errors "look good".

Above is an ARIMA(p, d, q) model.

Confused Yet?

Q: How do we account for seasonality, period s?

A: We do almost the exact same thing, except for period s:

- Look at $\{Y_t^*, Y_{t+s}^*, Y_{t+2s}^*, \ldots\}$. Are they stationary? If not, difference D times until they are.
- Guesstimate orders P and Q similarly to before.
- Fit "multiplicative ARMA(P, Q)" process, period s:

$$\left(Y_t^* - \Phi_1 Y_{t-s}^* - \cdots - \Phi_P Y_{t-Ps}^*\right) \phi(Y_t^*) = \left(Z_t + \Theta_1 Z_{t-s} + \cdots + \Theta_Q Z_{t-Qs}\right) \theta(Z_t)$$

- Repeat above 2 steps until all "looks good".

Above is an ARIMA$(p, d, q)(P, D, Q)_s$ process.

SAS Code

If you're still with me ...

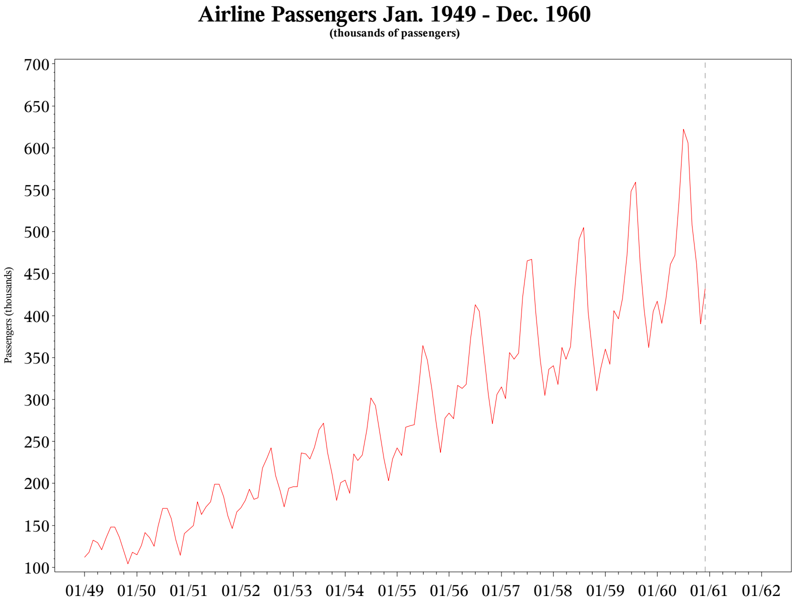
$$Y_t = \log(\texttt{pass}_t) \sim \text{ARIMA}(0,1,1) \times (0,1,1)_{12} :$$

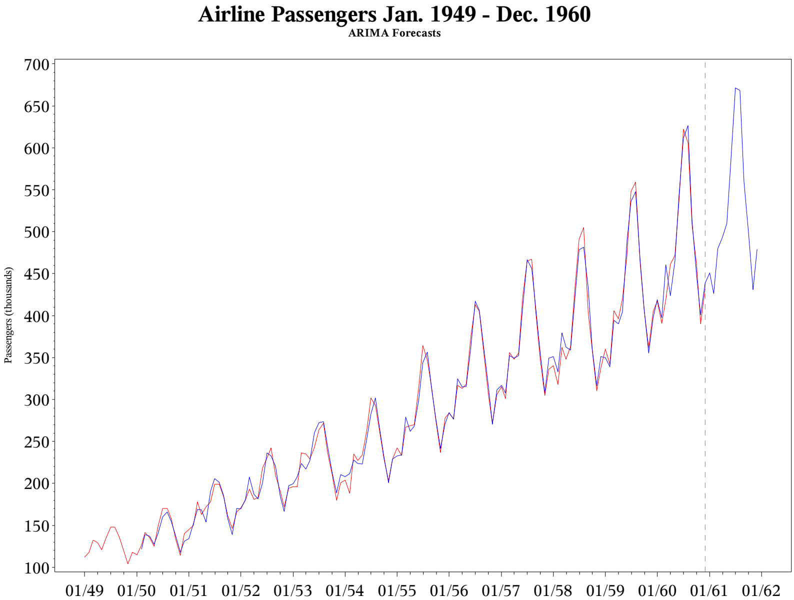
$$(Y_t - Y_{t-1})(Y_t - Y_{t-12}) = (Z_t - \theta_1 Z_{t-1})(Z_t - \Theta_1 Z_{t-12})$$

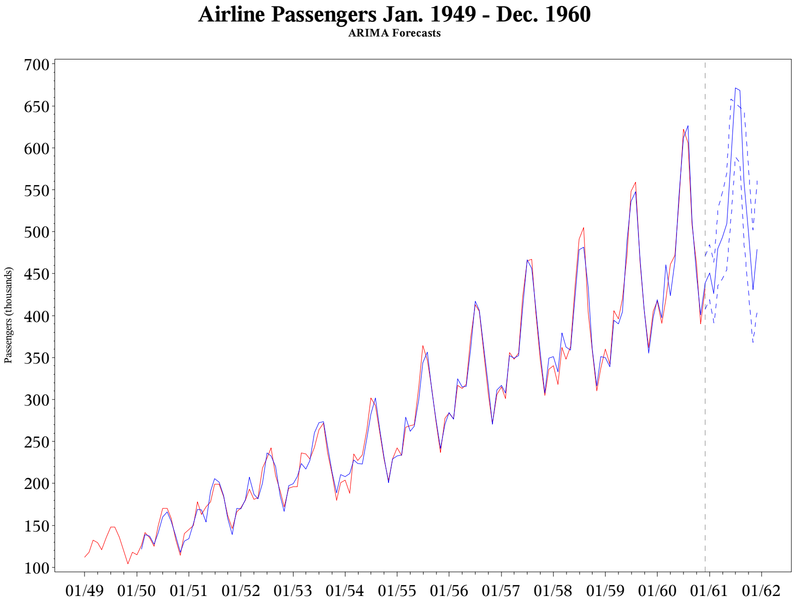
Forecasting 3 weeks ahead, ARIMA

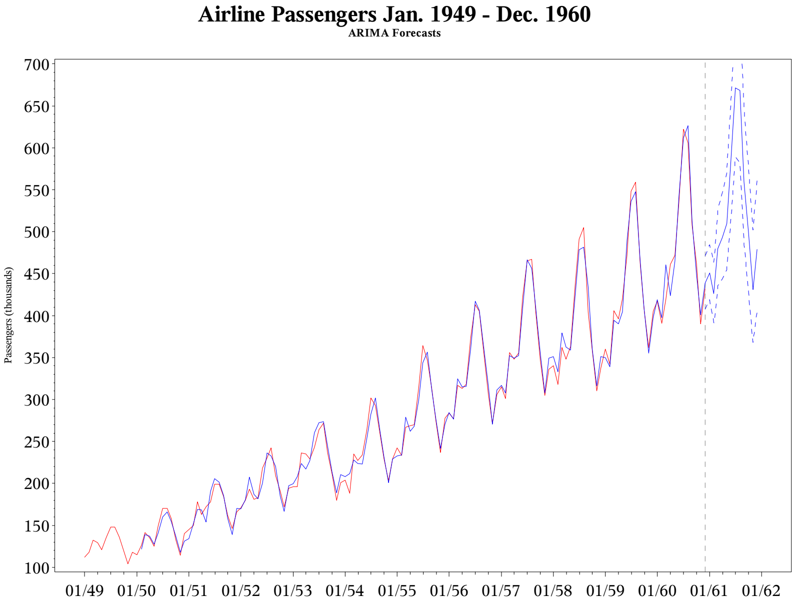
```
PROC ARIMA data=airline;
  IDENTIFY var=lpass( 1, 12 ) noprint;
  ESTIMATE q=( 1 )( 12 ) noint method=ML noprint;
  FORECAST lead=12 out=forearima id=date interval=month noprint;
RUN;
QUIT;
```

Compare with Moving Average


Airline Passengers Jan. 1949 - Dec. 1960
(thousands of passengers)
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62


Airline Passengers Jan. 1949 - Dec. 1960
ARIMA Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62


Airline Passengers Jan. 1949 - Dec. 1960
ARIMA Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62


Airline Passengers Jan. 1949 - Dec. 1960
ARIMA Forecasts
Passengers (thousands)
700
650
600
550
500
450
400
350
300
250
200
150
100
01/49
01/50
01/51
01/52
01/53
01/54
01/55
01/56
01/57
01/58
01/59
01/60
01/61
01/62

Beware the defaults!

SAS Code

```
symbol1 i=join c=red mode=include;
symbol2 i=join c=blue mode=include;
symbol3 i=join c=blue l=20 mode=include;

proc gplot data=forearima;
  plot pass*date=1
  forecast*date=2
  l95*date=3
  u95*date=3 / overlay ...;
run;
quit;
```

Which Method Should be Used?

We used three methods, would like to try others later.

Q: Which method should be used?

- Idea: The one that makes the best forecasts!
  - Make k-month-ahead forecasts for the last n months of the data.
    - For $i = 1, \ldots, n$, remove last i months of the data, then make forecasts for k months in the future.
  - For each method, compare forecasts to actuals.
  - Use forecasts from the method that made the most accurate forecasts.

How Do We Judge Forecasts?

- General standard: Mean Absolute Prediction Error (MAPE):

$$\text{MAPE} = 100 \times \sum_{t=1}^{T} \frac{|\text{forecast}_t - \text{actual}_t|}{\text{actual}_t},$$

  Gives average percentage off (zero is best!).

- Sometimes different methods best for different horizons.

How Do We Do This with SAS?

Easy way: Forecast Server or High Performance Forecasting!

- Follows (and generalizeds) our framework.
- Implements our methods.
- Allows us to add our own methods.

Harder (but cheaper) way: Program it ourselves.

How Do We Do This with SAS?

SAS Code Excerpt

```
DATA results;
  SET all;     *merged results, sorted by method;
  ape3 = 100*abs( pass - forecast3 )/pass;

PROC MEANS data=results noprint;
  BY method;
  VAR ape3;
  OUTPUT OUT=mapes MEAN( ape3 ) = mape3 / noinherit;

DATA mapes;
  SET mapes;
  IF method = 'arima' THEN CALL SYMPUT( 'mapearima', mape3 );
  IF method = 'exsm' THEN CALL SYMPUT( 'mapeexp', mape3 );
  IF method = 'mave' THEN CALL SYMPUT( 'mapemave', mape3 );

%LET mapev = &mapearima, &mapeexp, &mapemave;

DATA _null_;
  IF MIN( &mapev ) = &mapearima THEN CALL SYMPUT( 'best', 'arima' );
    ELSE IF MIN( &mapev ) = &mapeexp THEN CALL SYMPUT( 'best', 'exsm' );
    ELSE IF MIN( &mapev ) = &mapemave THEN CALL SYMPUT( 'best', 'mave' );

DATA bestforecasts;
  SET fore&best;
RUN;
```

Are Our Overall Forecasts Better?

- Better forecasts in training set no guarantee of better forecasts overall!
- Happily, we often *do* get better forecasts in general.

What's Next?

Multivariate Models!

- Takes account of holidays/other irregularities.
- Allows for scenario forecasting!

How will we do this?

How Will We Do This?

One solution: Multivariate ARIMA (transfer models):

$$Y_t = \beta_0 + \sum_{i=0}^{I} \beta_i X_{t-i} + Z_t, \quad Z_t = \text{ARIMA process}$$

- Works all right (using `PROC ARIMA`), but
- Very complicated to use,
- Results not very good/useful!

One big problem: Parameters are fixed over time.

- One outlier (e.g., Sept 11) could screw up entire model.
- If parameters could change over time, model would be (much) more flexible.

How Will We Do This?

Another solution: *State Space* (or *Hidden Markov*) Models

$$Y_t = \beta_{0t} + \sum_{i=0}^{I} \beta_{it} X_{t-i} + Z_t, \quad Z_t = \text{Normal process}$$

- Parameters change (slowly) over time.
  - Modeled by separate equation.
- Complicated, but flexibility makes it worth it.
- Problem: SAS doesn't implement it!
  - `PROC STATESPACE`: Nope! (misleading name)
  - `PROC UCM`: Closer, but still not there.
  - `PROC IML`: Can do it, but a fair bit of work.
  - (Almost) no one else (R, S+, SPSS) does, either.
  - My next research project!

Further Resources

- John C. Brocklebank and David A. Dickey.
  SAS for Forecasting Time Series.
  SAS Institute, 2003.
- Chris Chatfield.
  Time-Series Foreasting.
  Chapman and Hall, 2000.

Nate Derby: `http://nderby.org`

`nderby@sprodata.com`