Validating a Forecast

P481 Supplement to Chapter 7

What do we want to get out of a forecasting model? Would we just like to know that the model fits with our existing data or do we hope that it can predict the future? How can we compare two or more forecasting models in a fair way?

Example

A small food manufacturer sells gourmet products, and has recently introduced a new kind of tomato sauce. Managers have tracked the demand for cases of this sauce for the past 8 weeks as follows:

Week	Demand
1	134
2	205
3	218
4	180
5	206
6	250
7	203
8	228

Alice thinks there is a trend in the data and wants to use a **static forecast**. So she runs a linear regression and finds that LEVEL=162.8, TREND=9.2. Computing LEVEL + 9 * TREND = 244, Alice forecasts demand for 244 cases in week 9.

Bob wants to use a **moving average of 4 periods**. So he finds AVERAGE(206, 250, 203, 228) = 222. Bob forecasts demand for 222 cases in week 9.

These two forecasts differ by quite a lot—Alice's forecast is 10% higher than Bob's. Which one should the company use for its planning purposes?

Hold out sample

The company would like to use the more accurate forecasting method, but until the demand from week 9 is realized, we won't know whose forecast was better. Even after the demand from week 9 is realized, we still won't be able to tell for sure because some of the change from week 8 to week 9 is noise.

Using a hold out sample is a more systematic way to validate and compare forecasting methods. The idea is to hold out part of the data to use for validation. In this way, we simulate "new" data being observed. Let's see how this applies to the example.

Alice and Bob agree to use the first 5 weeks as the **training set**, to create the model, and the last 3 months as the **hold out sample** or **validation set**, to test the model's performance.

Results

Alice re-runs her regression on just the first 5 weeks of data (it would be cheating if she could "see into the future" and use the hold out data!). She gets a new estimate of LEVEL=152.9, TREND = 11.9. Both Alice and Bob make their forecasts for weeks 6-8:

Week	Demand	Alice	Bob
1	134	Forecast	Forecast
2	205		
3	218		
4	180		
5	206		
6	250	224.3	202.3
7	203	236.2	213.5
8	228	248.1	209.8

The hold out sample is the data points 250, 203, and 228 corresponding to weeks 6-8.

Try This

Using the formulas learned in class, calculate the Mean Squared Error (which measures the size of the error) and the Bias (which measures the direction of the error) for both Alice's and Bob's forecasts.

Study questions

1. Which forecast seems to perform better? Why?
2. Is this a fair comparison of the two forecasting methods?
3. How would the size of the hold out sample matter?