**Primer on Classical Statistical Forecasting**

**SARIMAX Modeling:**

Auto-Regressive (AR) Models:

In an auto-regressive model, the variable of interest at time is forecast as a linear combination of past values of the variable. The general structure has the following form: [1,2]

In the above equation, is the order of the model, or the number of *lagged values* of , is the intercept, and is the error at time .

We can also introduce another common notation, , the *lag operator*: [3]

From the lag operator, we can define the *lag polynomial for AR processes*, , also written as . In this notation, we can rewrite the first equation (ignoring the intercept) as: [4,5]

We can rearrange this to get the first AR equation:

Moving Average Models:

In a moving average model, the variable of interest at time is forecast as a linear combination of past forecast errors, . The general structure has the following form: [2, 6]

In the above equation, is the order of the model, or the number of *past errors* being used, is the intercept, and is the error at time .

With the *lag polynomial for MA processes* notation, our MA model becomes: [4]

ARMA Model:

We can then combine them to create an *Auto-Regressive Moving Average* (ARMA) model: [2,4]

Or, using our other notation, dropping the constant, and rearranging a bit, we get:

Differencing*:*

In general, it is easier to forecast from a *stationary* series, one whose statistical properties do not depend on the time at which the series is observed. We can make a series stationary by performing *differencing*, or taking the difference between two consecutive values of the series: [7]

Using lag notation, we can write this as:

We use several different stationarity tests, such as the *Augmented Dickey-Fuller Test*, to check for stationarity and perform the necessary number of differencings to achieve stationarity. The number of differencings is known as parameter . We can also write the relationship as follows: [2]

ARIMA:

Now, we can combine each of these terms to create an *Auto-Regressive Integrated Moving Average* (ARIMA) model: [2]

Using our lag operator notation and dropping the constant, this is: [5]

Rearranging them, we arrive at the common form: [5]

This ARIMA model is often written in shorthand as .

ARIMAX:

However, this model only used errors and past values of the series to forecast future values. If we wish to include more variables, , named *exogenous* variables, we now have an *Auto-Regressive Integrated Moving Average with eXogenous variables* (ARIMAX) model. To make this change, we need only add a linear combination of the exogenous variables to our original ARIMA model: [2]

Using our lag operator notation and dropping the constant, this is: [5]

Or: [5]

Seasonality:

Often, our data has a *seasonal* component, or a cycle that repeats every certain number of time periods. Perhaps our data shows significant changes throughout the month, or week, day. In that case, we need to consider the period of the season and integrate it into our model. In this case, our variable of interest might depend not just several past instances or errors, but on values from a past seasonal period. Fortunately, we treat the seasonal component with an ARIMA model, as well. Though, we now use the parameters , where is the *seasonal auto-regressive* parameter, is the *seasonal differencing* parameter, is the *seasonal moving average* parameter, and is the *seasonality*, or the number of time periods in a season. For example, on daily data that has a weekly seasonal component, .

We augment our model to become a model with the form: [2]

As above, we can rearrange this and drop the constant to arrive at a more common form: [5, 9]

**EST Modeling:**

Simple Exponential Smoothing:

In *Simple Exponential Smoothing*, the estimate of the value of a time series is modeled as a weighted average of the most recent values that came before it: [9]

Equivalently, this can be thought of as a combination of the most recent value, , and the previous forecast, : [9]

Finally, to aid us in writing further exponential smoothing models, we will write this in *component form* to see that the equation has two pieces, the *Forecast* and the *Level Smoothing*: [9]

Here, the term is the estimate of the level of the series at time and is the smoothing parameter for the level.

Trended Method:

If the time series displays a trend, we will add another component, a *Trend Smoothing* one. This model is called the *Holt’s Linear Trend*: [10]

Here, the term is the estimate of the level of the series at time , is the smoothing parameter for the level. is the estimate of the trend, or slope, of the series at time , and is the trend smoothing parameter.

However, it is often the case that a trended time series is expected to approach a constant. If that is the case, we can introduce a *dampening* constant , which serves to force the series to approach a constant as time goes on. [10]

Seasonality

If our time series has a seasonal component, we will add a *Seasonal Smoothing* one to our model. This has one of two forms, *additive* or *multiplicative*. We will show two examples, the additive one: [11]

and the multiplicative one: [11]

In these equations is the seasonality, , and .

For brevity, we will stop here, as we have shown how one builds increasingly complex *Exponential Smoothing* models, but we are still missing a significant piece, the errors, .

Just as with the multiplicative component, we can consider the errors in an additive or multiplicative way. We can also introduce a dampening parameter into either of the two above models. From this, the number of possible models grows quite large, so we will introduce the following notation: [12]

The is the *error* term, the is the *trend* term, and the is the *seasonality* term. The errors can be additive or multiplicative, so, . The trend can be none (absent), additive, or additive and dampened, so, . The season can be none, additive, or multiplicative, so, . Note, we are not including multiplicative trend models, as they tend to do poorly. Simple arithmetic shows that, with these simple parameters, we have a taxonomy of 18 models. We will provide a table for them and close this primer. [12]

A picture containing text, receipt, screenshot

Description automatically generated