

# Nov 27 Notes

*Hongdou Li*

*11/27/2018*

## Exponential Smoothing (ES) aka Holt-Winters Methodology

The objective is to predict  $Y_{n+h}$  given the history of observations  $\{y_1, y_2, \dots, y_n\}$ . Using exponential smoothing techniques we do so with a set of recursive equations that do not require distributional assumptions.

We will use different sets of equations to fit data depending on whether the data has

- (1) no trend + no seasonality <- single/simple ES
- (2) trend + no seasonality <- Double ES
- (3) trend + seasonality <- Triple ES

## Single Exponential Smoothing (SES)

“Level” Equation:  $a_t = \alpha y_t + (1 - \alpha)a_{t-1}$

Forecast Equation:  $\hat{y}_{t+h} = a_t$  for  $h=1,2,3,\dots$

\*  $0 \leq \alpha \leq 1$  is a **smoothing parameter**. small  $\alpha$  performs more smoothing and large  $\alpha$  performs less. Note, if  $\alpha = 0$ ,  $a_t = a_0$  (super smooth) and if  $\alpha = 1$ ,  $a_t = y_t$  (no smoothing).

\*  $\alpha = 0.2$  is typically a good choice, but an optimal value of  $\alpha$  can be determined using the data. In particular  $\hat{\alpha}$  is the value that minimizes squared error loss function:

$$\sum_{t=1}^n e_t^2 = \sum_{t=1}^n (y_t - \hat{y}_t)^2$$

This is often referred to as an Exponentially Weighted Moving Average(EWMA)

This is why:

$$\begin{aligned} a_t &= \alpha y_t + (1 - \alpha)a_{t-1} \\ &= \alpha y_t + (1 - \alpha)(\alpha y_{t-1} + (1 - \alpha)a_{t-2}) \\ &= \alpha y_t + (1 - \alpha)\alpha y_{t-1} + (1 - \alpha)^2 a_{t-2} \\ &= \alpha y_t + (1 - \alpha)\alpha y_{t-1} + (1 - \alpha)^2(\alpha y_{t-2} + (1 - \alpha)a_{t-3}) \\ &= \alpha y_t + (1 - \alpha)\alpha y_{t-1} + (1 - \alpha)^2 \alpha y_{t-2} + (1 - \alpha)^3 a_{t-3} \\ &\quad \cdot \\ &\quad \cdot \\ &\quad \cdot \\ &= \alpha \sum_{i=0}^{t-1} (1 - \alpha)^i y_{t-i} + (1 - \alpha)^t a_0 \end{aligned}$$

where  $a_0$  is the initial value of the recursion. A common choice for  $a_0$  is  $\bar{y}$  but as we see here, what we choose does not really matter.

## Double Exponential Smoothing (DES)

“Level” Equation:  $a_t = \alpha y_t + (1 - \alpha)(a_{t-1} + b_{t-1})$  (Weighted average of the observed response at time  $t$  and the model’s corresponding prediction of that observation.)

“Trend” Equation:  $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$  (weighted average of current trend component and historical prediction of that same component.)

“Forecast” Equation:  $\hat{y}_{t+h} = a_t + hb_t$  for  $h=1,2,3,\dots$

where  $0 \leq \alpha \leq 1$  and  $0 \leq \beta \leq 1$  are smoothing parameters, where small values provide more smoothing a large values provide less.

## Triple Exponential Smoothing (TES)

“Level” Equation:  $a_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(a_{t-1} + b_{t-1})$  (weighted average of seasonally adjusted observation and a non-seasonal forecast of it)

“Trend” Equation:  $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$  (weighted average between the current seasonal index and the seasonal index of the same period in the previous season)

“Seasonal” Equation:  $s_t = \gamma(y_t - a_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$

Forecast Function :  $\hat{y}_{t+h} = a_t + hb_t + s_{t+h-m}$  for  $h=1,2,3,\dots$

where  $0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, 0 \leq \gamma \leq 1$  are smoothing parameters that behave as usual, and that can be estimated from the observed data.

This formulation is referred to as “additive” but when heteroscedasticity is present, we may opt to use a multiplicative version.

LE •  $a_t = \alpha\left(\frac{y_t}{s_{t-m}}\right) + (1 - \alpha)(a_{t-1} + b_{t-1})$

TE •  $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$

SE •  $s_t = \gamma\left(\frac{y_t}{a_{t-1} + b_{t-1}}\right) + (1 - \gamma)s_{t-m}$

FE •  $\hat{y}_{t+h} = (a_t + hb_t)s_{t-m}$  for  $h=1,2,3,\dots$