

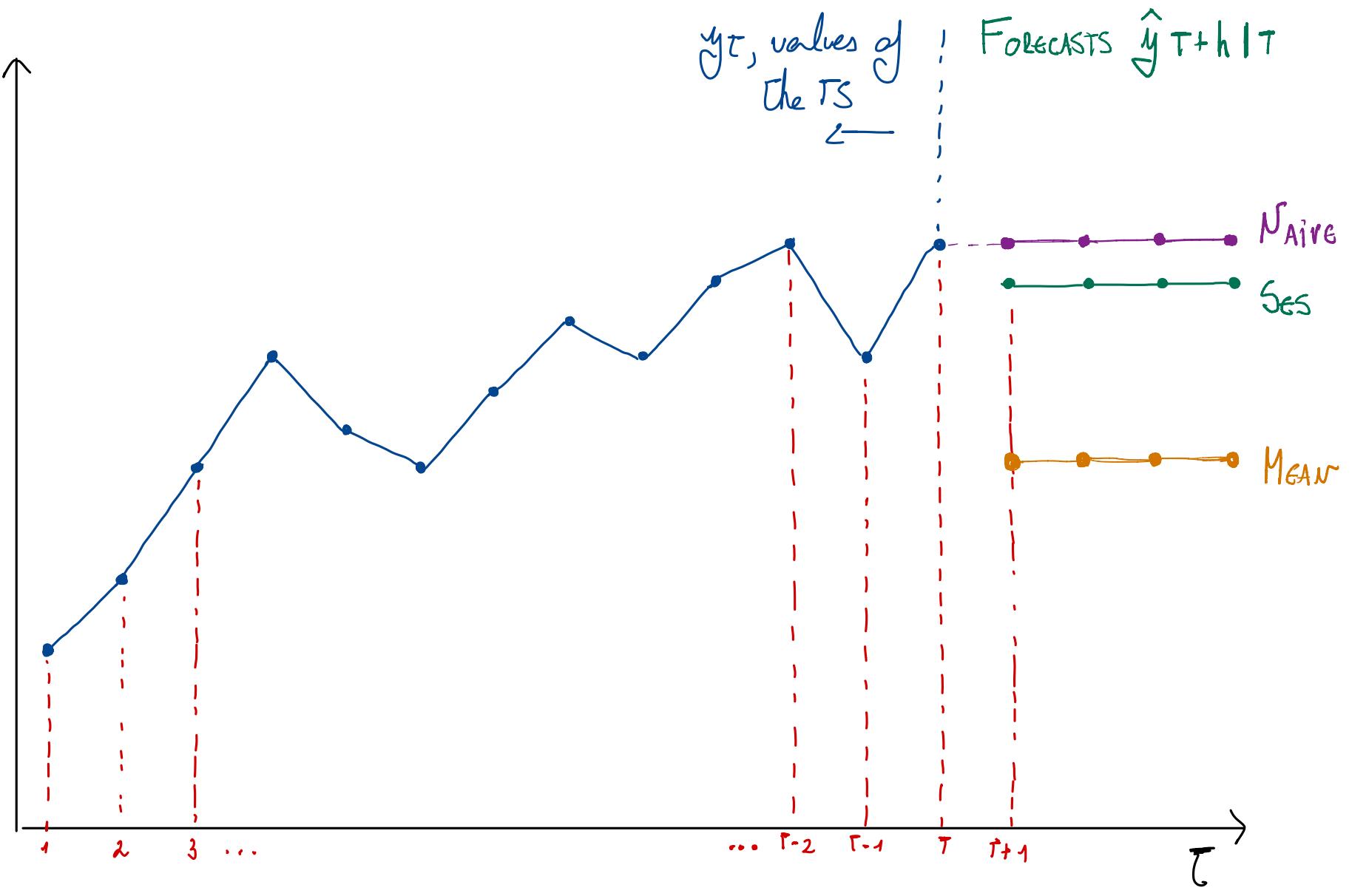
Simple Exponential Smoothing (SES)

Prof. Juan Garbayo

Time Series Analysis

CONTENTS

- ① SES vs Naive vs Mean
- ② Geometric progression - basic properties
- ③ SES - weighted average form
- ④ SES - components form - the level
- ⑤ Filled values and flat forecasts
- ⑥ Exponential Moving Average vs Simple Moving Average



ALL WEIGHT TO
LAST OBSERVATION

SAME WEIGHT TO ALL OBSERVATIONS, NO MATTER HOW DISTANT IN THE PAST

EXPONENTIAL DECREASE

... SES WEIGHTS

More weight to recent observations
EXPONENTIALLY DECAYING WEIGHTS

Geometric sequence (1/3)

$$r = \frac{a_n}{a_{n-1}} \rightarrow \text{ratio (given)} \rightarrow \text{Definition}$$

a_1 → given

$$a_2 = r \cdot a_1$$

$$a_3 = r^2 \cdot a_1$$

⋮

$$a_N = r^{N-1} \cdot a_1$$

$$\text{If } r < 1 \Rightarrow \lim_{n \rightarrow \infty} a_n = 0$$

A specific example:

$$a_1 = \alpha \quad \alpha \in (0, 1]$$

$$r = (1 - \alpha) \Rightarrow (1 - \alpha) \in [0, 1)$$

$$a_n = \underbrace{\alpha}_{< 1} (1 - \alpha)^{n-1}$$

a_n decreases exponentially with $\uparrow n$

$$\lim_{n \rightarrow \infty} a_n = 0$$

Geometric sequence (2/3)

SUM OF THE FIRST n TERMS OF A GEOMETRIC SEQUENCE

$$S_n = a_1 + \underbrace{a_1 \cdot r}_{a_2} + \underbrace{a_1 \cdot r^2}_{a_3} + \dots + \underbrace{a_1 \cdot r^{n-1}}_{a_n}$$

$$r \cdot S_n = a_1 \cdot r + a_1 \cdot r^2 + a_1 \cdot r^3 + \dots + a_1 \cdot r^n$$

$$S_n - r \cdot S_n = S_n(1-r) = a_1 \cdot a_1 \cdot r^n = a_1 (1-r^n)$$

For $r < 1$, S_n converges

Solving for $S_n \rightarrow S_n = \frac{a_1(1-r^n)}{(1-r)} \rightarrow$ If $r < 1 \Rightarrow \lim_{n \rightarrow \infty} S_n = \frac{a_1}{1-r}$

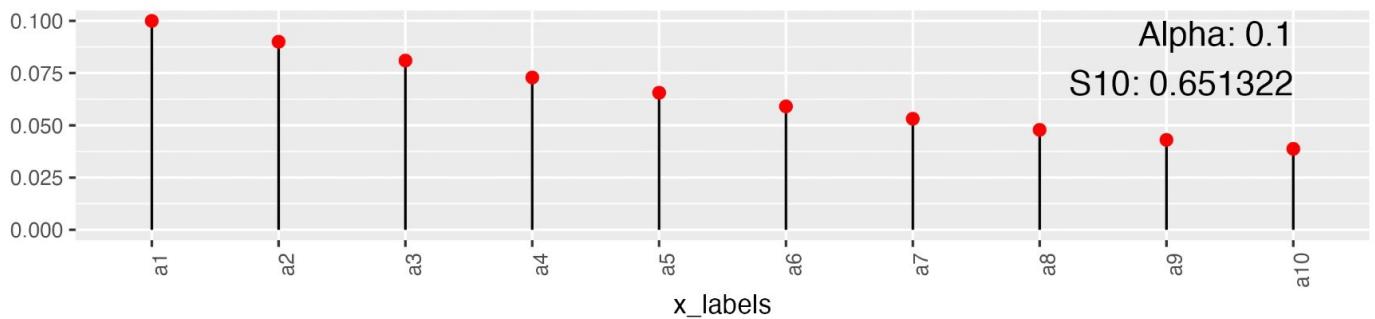
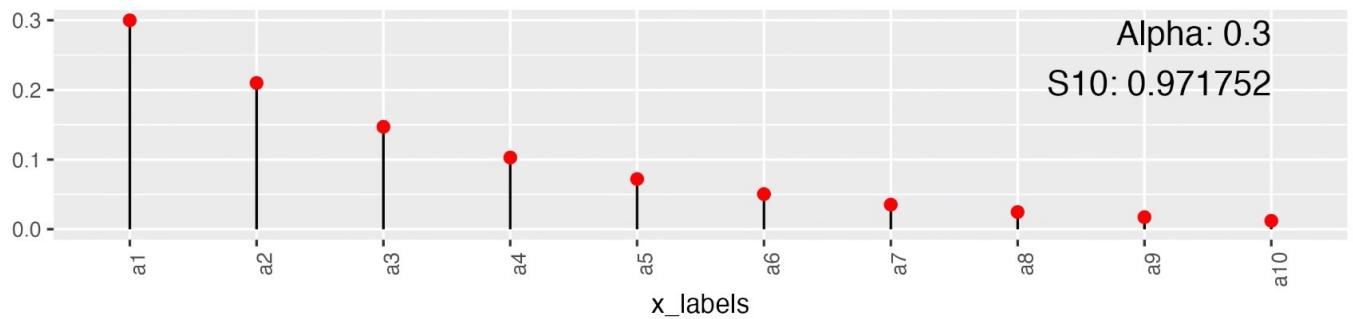
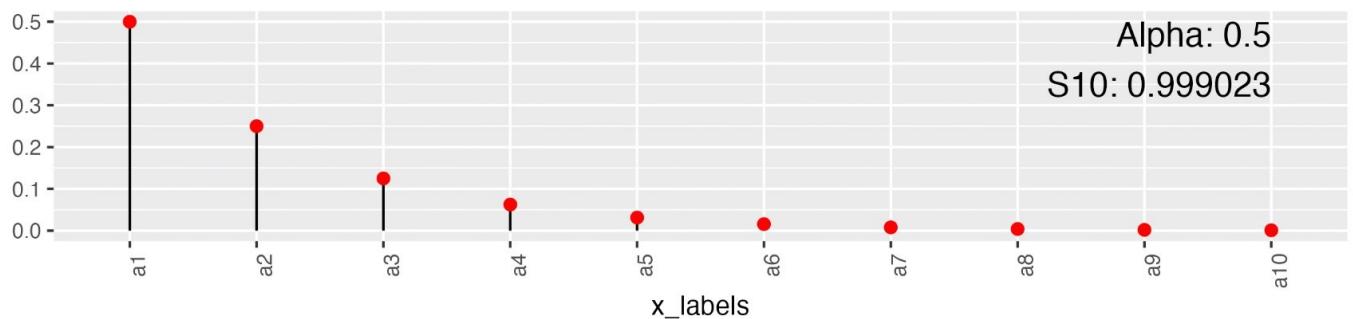
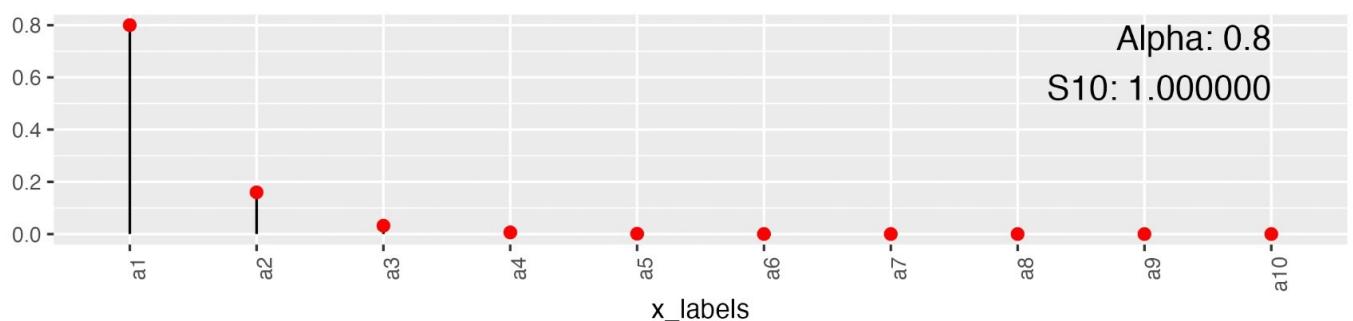
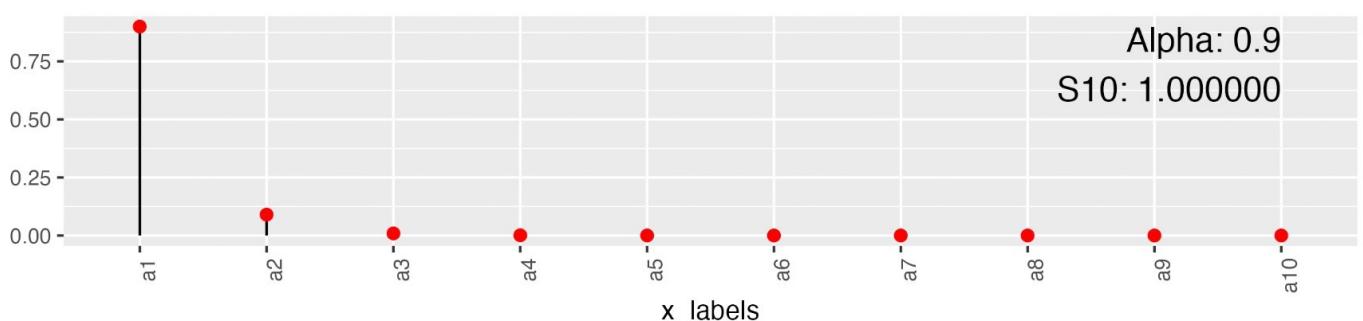
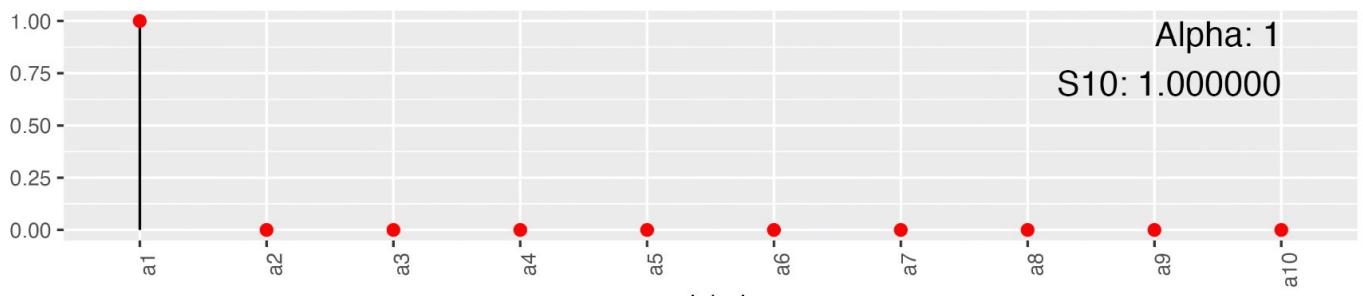
Specific case: $a_1 = \alpha \in (0, 1]$ $\left. \begin{array}{l} \\ r = (1-\alpha) < 1 \end{array} \right\} \rightarrow S_n = \frac{\alpha(1-(1-\alpha)^n)}{1-(1-\alpha)} = 1 - (1-\alpha)^n$

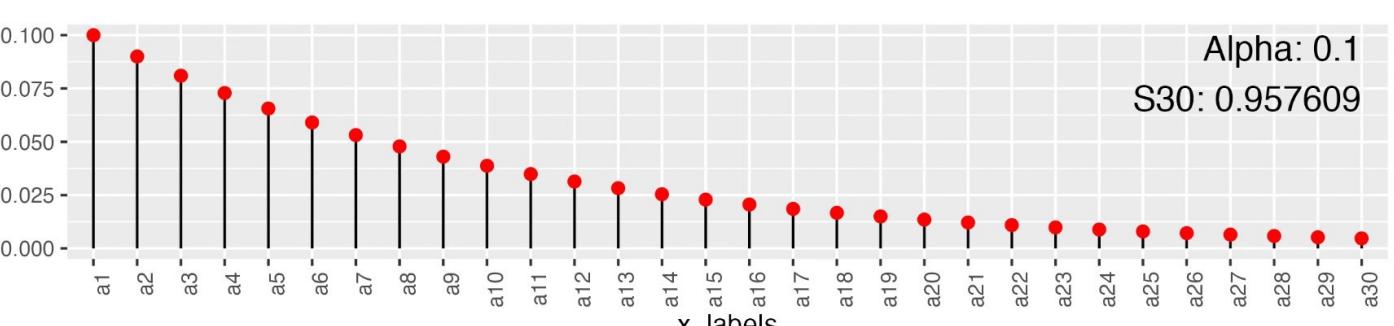
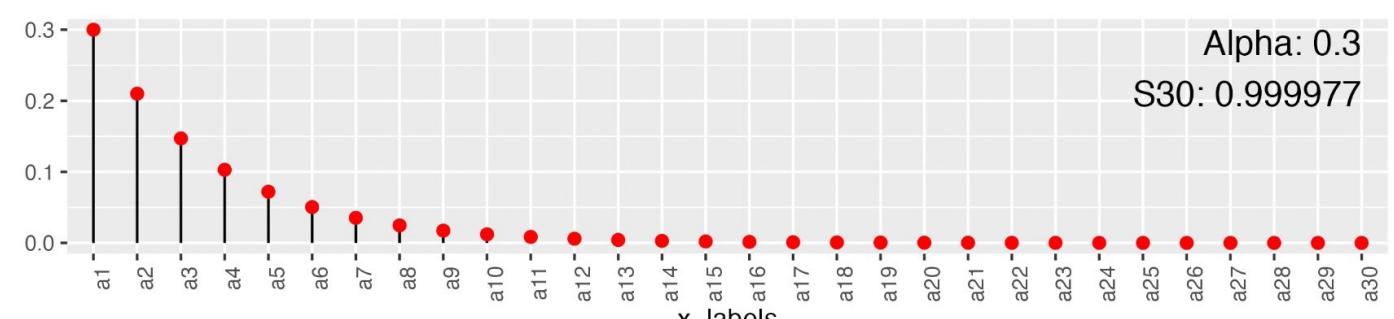
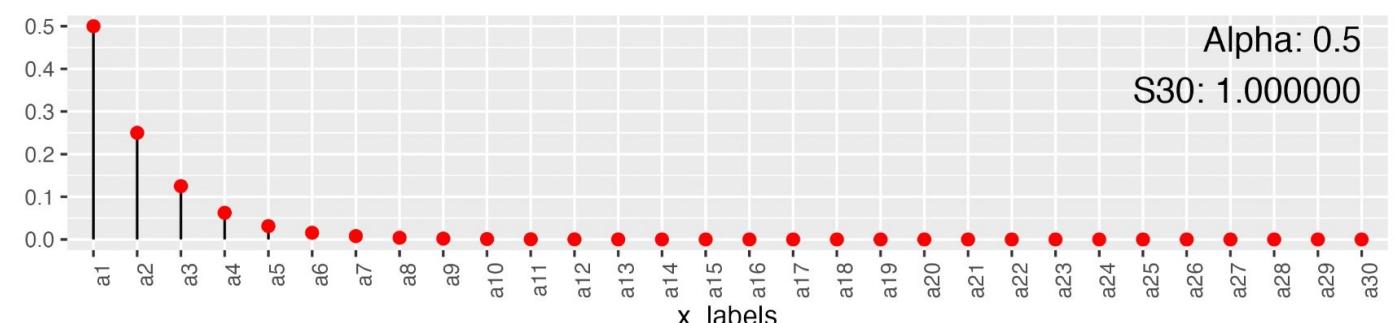
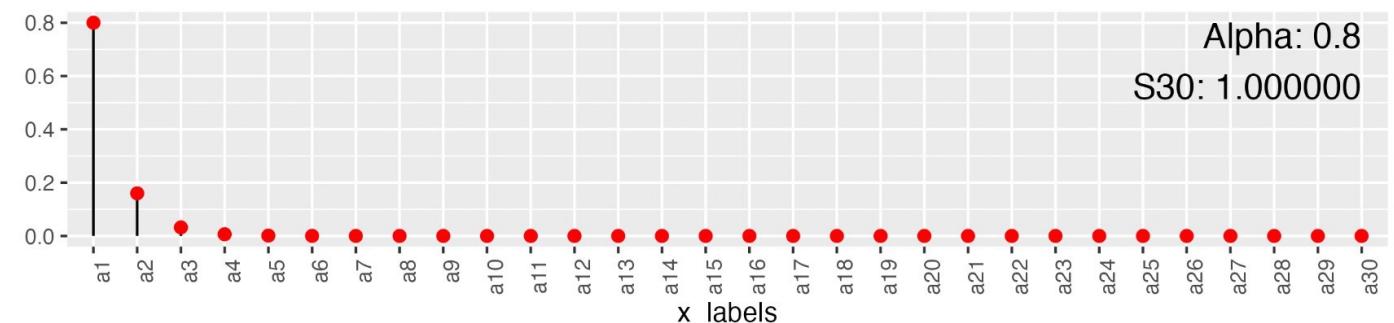
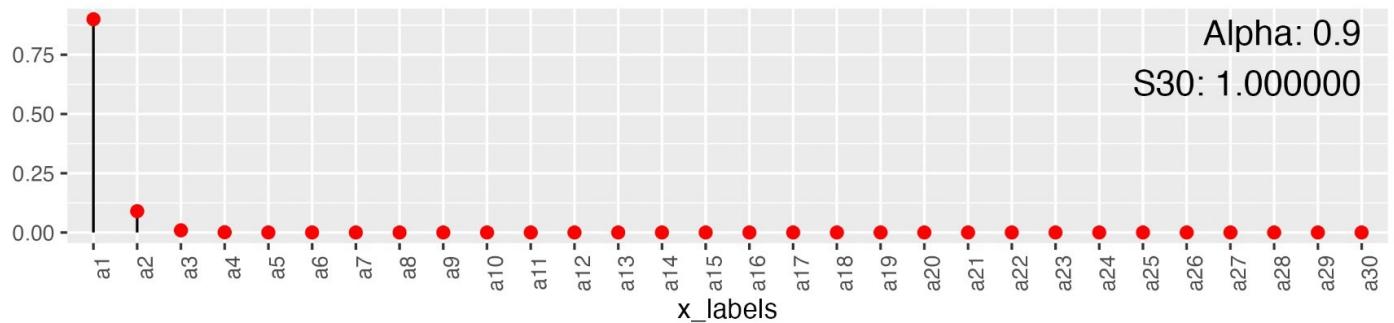
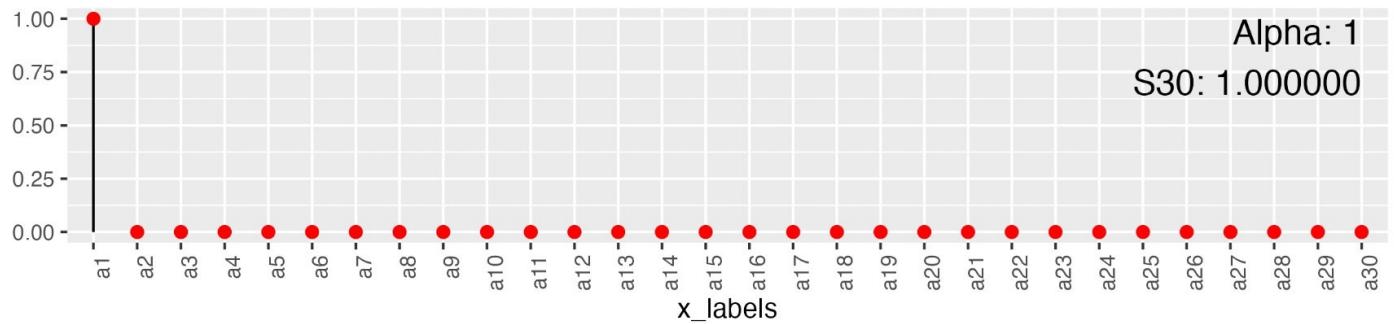
$$\left[S_n = (1 - (1-\alpha)^n) \text{ ; } \lim_{n \rightarrow \infty} S_n = 1 \right]$$

Geometric sequence (3/3)

We have shown that, for $a_1 = \alpha \in [0, 1]$ |
 $r = (1-\alpha) \in [0, 1)$

- $a_n = \alpha \left(\frac{1}{1-\alpha}\right)^{n-1}$ ~ EXPONENTIAL DECAY
 $\lim_{n \rightarrow \infty} a_n = 0$ DISTINCT OBSERVATIONS HAVE NEGLIGIBLE WEIGHTS
- $S_n = 1 - (1-\alpha)^n$; $\lim_{n \rightarrow \infty} S_n = 1$ α CAN BE INTERPRETED AS THE PERCENT WEIGHT ASSIGNED TO y_i





WEIGHTED AVERAGE FORM

- Given these properties, we may produce a forecast with exp decreasing weights:

$$\hat{y}_{T+1|T} = \underbrace{\alpha_1}_{\alpha_1} y_T + \underbrace{\alpha(1-\alpha)}_{\alpha_2} y_{T-1} + \underbrace{\alpha(1-\alpha)^2}_{\alpha_3} y_{T-2} + \dots + \underbrace{\alpha(1-\alpha)^{T-1}}_{\alpha_T} y_1$$

$\alpha_n = \alpha(1-\alpha)^{n-1}$

$1: T-(T-1)$

- We may express this in a compact manner using the sum operator:

$$\begin{aligned}\hat{y}_{T+1|T} &= \sum_{i=1}^T \alpha_i y_{T-(i-1)} = \sum_{i=1}^T \alpha(1-\alpha)^{i-1} y_{T-(i-1)} = \\ &= \sum_{j=0}^{T-1} \alpha(1-\alpha)^j y_{T-j}\end{aligned}$$

Re-indexing: $j = i-1$

WEIGHTED AVERAGE FORM

$$\hat{y}_{T+h|T} = \sum_{j=0}^{T-1} \alpha(1-\alpha)^j y_{T-j}$$

By definition of the SES Model

$$\hat{y}_{T+h|T} = \hat{y}_{T+1|T}$$

FLAT FORECASTS

FITTED VALUES USING W. Av. Form

Fitted values are DEFINED as $\hat{y}_{t+1|t}$

$$\hat{y}_{T|T-1} = \alpha y_{T-1} + \alpha(1-\alpha) y_{T-2} + \alpha(1-\alpha)^2 y_{T-3} + \dots + \alpha(1-\alpha)^{T-2} y_1 + \alpha(1-\alpha)^{T-1} y_0$$

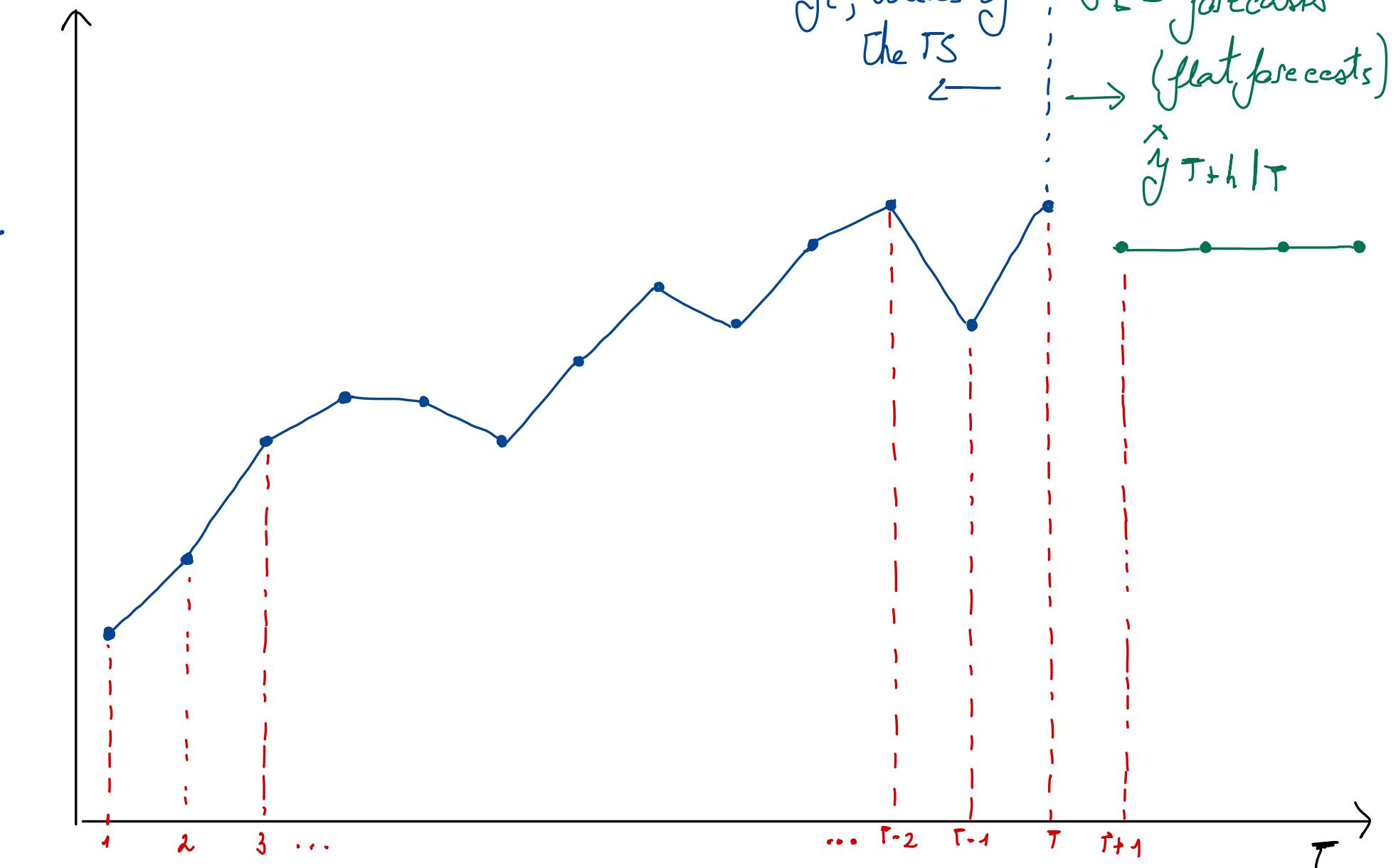
$$\hat{y}_{T-1|T-2} = \alpha y_{T-2} + \alpha(1-\alpha) y_{T-3} + \alpha(1-\alpha)^2 y_{T-4} + \dots + \alpha(1-\alpha)^{T-3} y_1 + \alpha(1-\alpha)^{T-2} y_0$$

$$\hat{y}_{3|2} = \alpha y_2 + \alpha(1-\alpha) y_1$$

$$\hat{y}_{2|1} = \alpha y_1$$

$$\hat{y}_{1|0} = ?$$

$\alpha \cdot y_0 ??$



PROBLEM:

- COMPUTATIONALLY QUITE INEFFICIENT → The amount of multiplications and sums may be substantially reduced
- NO FITTED VALUE AT $t=1$
- Only one parameter α to adjust the fit of our model

RECURSIVE EQUATION FOR THE FITTED VALUES

WEIGHTED AVERAGE FORM

$$\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha(1-\alpha)^j y_{T-j}$$

We are going to use an alternative procedure to compute the fitted values that will solve the problems outlined and enable us to interpret exponential smoothing in terms of "components". This will be most clear in trended and seasonal exponential smoothing.

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) \hat{y}_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \dots + \alpha(1-\alpha)^{T-1} y_1$$

(*)

$\hat{y}_{T+1|T}$ is a weighted average of y_T and \hat{y}_{T-1} . The weights are α and $\alpha(1-\alpha)$.

$$\hat{y}_{T|T-1} = \alpha y_{T-1} + \alpha(1-\alpha) \hat{y}_{T-2} + \alpha(1-\alpha)^2 y_{T-3} + \dots + \alpha(1-\alpha)^{T-2} y_1$$

$\hat{y}_{T|T-1}$ is a weighted average of y_{T-1} and \hat{y}_{T-2} . The weights are α and $\alpha(1-\alpha)$.

$$(1-\alpha) \hat{y}_{T|T-1} = \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \alpha(1-\alpha)^3 y_{T-3} + \dots + \alpha(1-\alpha)^{T-1} y_1$$

$(1-\alpha) \hat{y}_{T|T-1}$ is a weighted average of y_{T-1} and \hat{y}_{T-2} . The weights are $(1-\alpha)$ and α .

$$\hat{y}_{T+1|T} = \alpha \cdot y_T + (*) = \alpha \cdot y_T + (1-\alpha) \hat{y}_{T|T-1} \Rightarrow \hat{y}_{T+1|T} = \alpha y_T + (1-\alpha) \hat{y}_{T|T-1}$$

Recursive equation for the fitted values
No new info. Direct consequence of the weighted average form

RECURSIVE COMPUTATION OF FITTED VALUES

$$\hat{y}_{T+1|T} = \alpha y_T + (1-\alpha) \hat{y}_{T|T-1}$$

EQUATION AT $T = T+1$

C.V. $T = T+1 \Rightarrow T = T-1$

$$\hat{y}_{T|T-1} = \alpha y_{T-1} + (1-\alpha) \hat{y}_{T-1|T-2}$$

EQUATION AT GENERIC INSTANT T

NOTE THAT $\alpha + (1-\alpha) = 1$

The fitted value at T is a weighted average of:

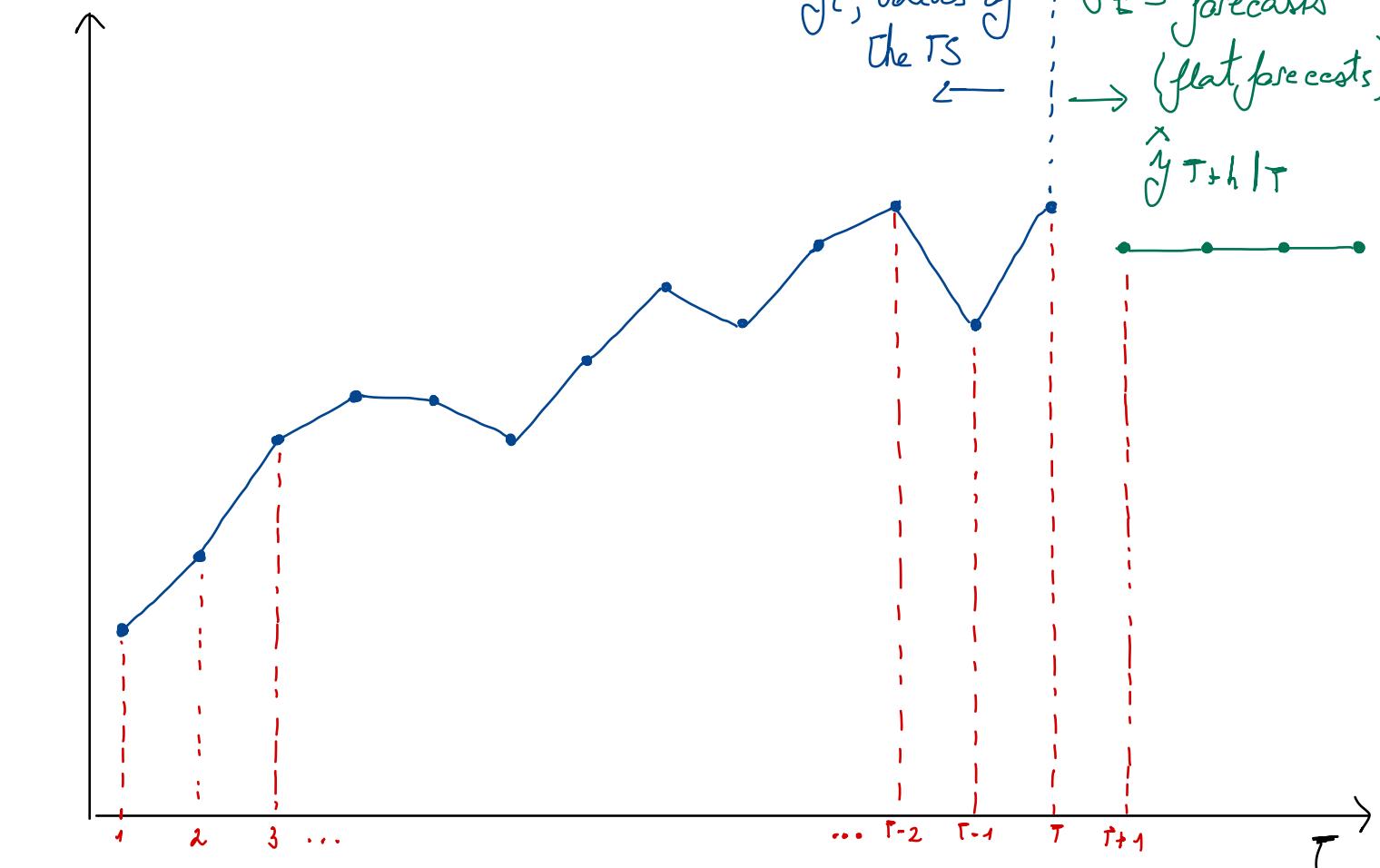
→ The observation at $T-1$ (y_{T-1})

→ The fitted value at $T-1$ ($\hat{y}_{T-1|T-2}$)

$$\begin{array}{rcl} T & | & \hat{y}_{T|T-1} \\ \hline T & | & \hat{y}_{T|T-1} = \alpha y_{T-1} + (1-\alpha) \hat{y}_{T-1|T-2} \\ T-1 & | & \hat{y}_{T-1|T-2} = \alpha y_{T-2} + (1-\alpha) \hat{y}_{T-2|T-3} \\ T-2 & | & \hat{y}_{T-2|T-3} = \alpha y_{T-3} + (1-\alpha) \hat{y}_{T-3|T-4} \\ \vdots & | & \vdots \\ 3 & | & \hat{y}_{3|2} = \alpha y_2 + (1-\alpha) \hat{y}_{2|1} \\ 2 & | & \hat{y}_{2|1} = \alpha y_1 + (1-\alpha) \hat{y}_{1|0} \\ 1 & | & \hat{y}_{1|0} = l_0 \end{array}$$

Initial value for the level

↳ NEW PARAMETER



FITTED VALUES AS UPDATING APPROXIMATION

Note that $\alpha + (1-\alpha) = 1$

$$\hat{y}_{T|T-1} = \alpha \hat{y}_{T-1} + (1-\alpha) \hat{y}_{T-1|T-2}$$

$$\hat{y}_{1|0} = l_0$$

EQUATION AT GENERIC INSTANT T

INITIAL LEVEL. GUESS. PARAMETER OF PROCESS.

$$\begin{aligned} & \hat{y}_{1|0} = l_0 \\ & \text{OBSERVATION: NEW INFO} \\ & \text{UPDATED APPROXIMATION} \quad \hat{y}_{2|1} = \alpha \hat{y}_1 + (1-\alpha) \hat{y}_{1|0} \\ & \hat{y}_{3|2} = \alpha \cdot y_2 + (1-\alpha) \hat{y}_{2|1} \\ & \vdots \quad \vdots \\ & \hat{y}_{T-2|T-1} = \alpha \cdot y_{T-1} + (1-\alpha) \hat{y}_{T-3|T-4} \\ & \hat{y}_{T-1|T-2} = \alpha \cdot y_{T-2} + (1-\alpha) \hat{y}_{T-2|T-1} \\ & \hat{y}_{T|T-1} = \alpha \cdot y_{T-1} + (1-\alpha) \hat{y}_{T-1|T-2} \end{aligned}$$

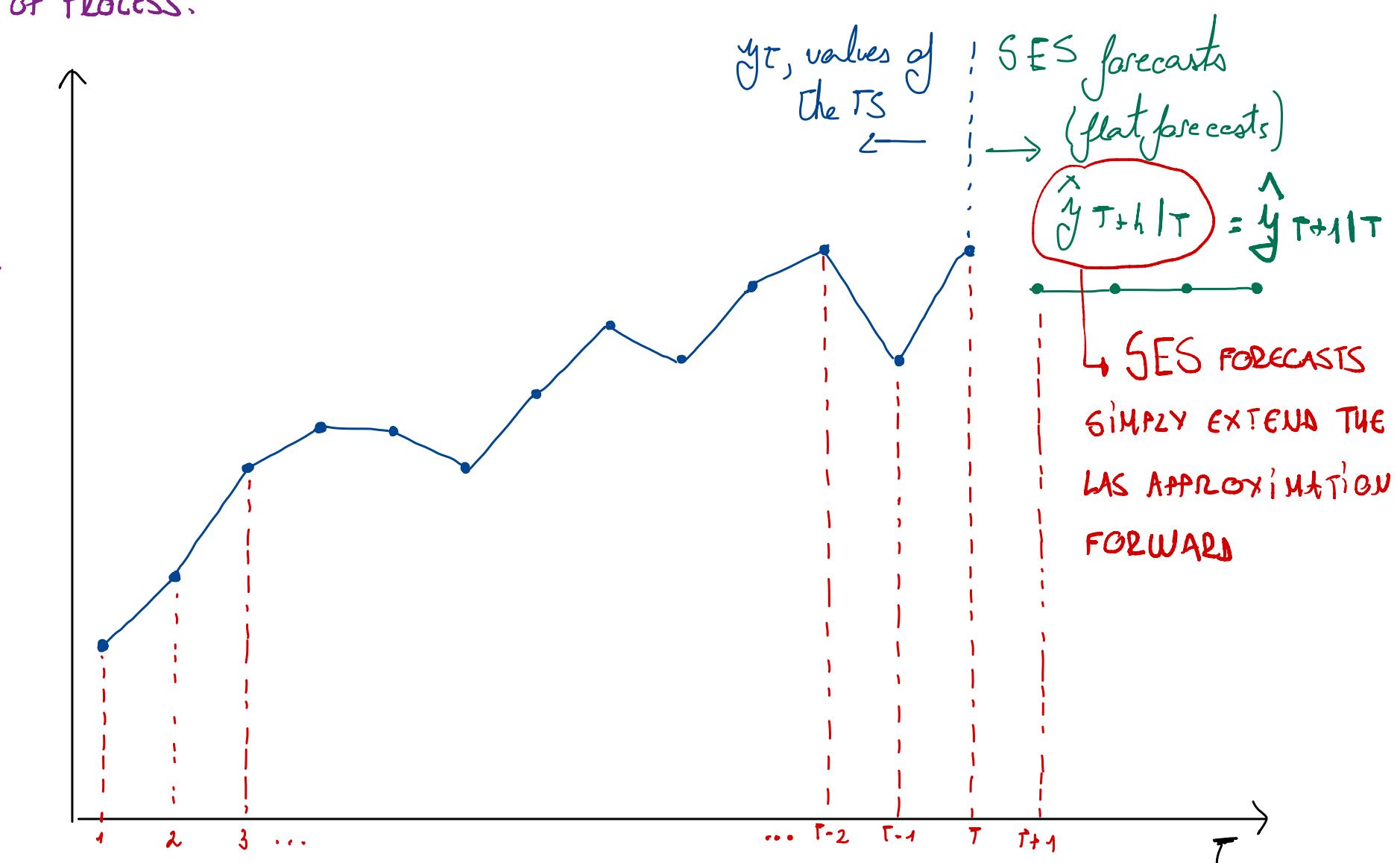
The fitted value at T $\hat{y}_{T|T-1}$ is a weighted average of:

- The observation at $T-1$ y_{T-1}
- The fitted value at $T-1$ $\hat{y}_{T-1|T-2}$

Getting a new observation at T y_T allows us to update our approximations for the next point in time $T+1$

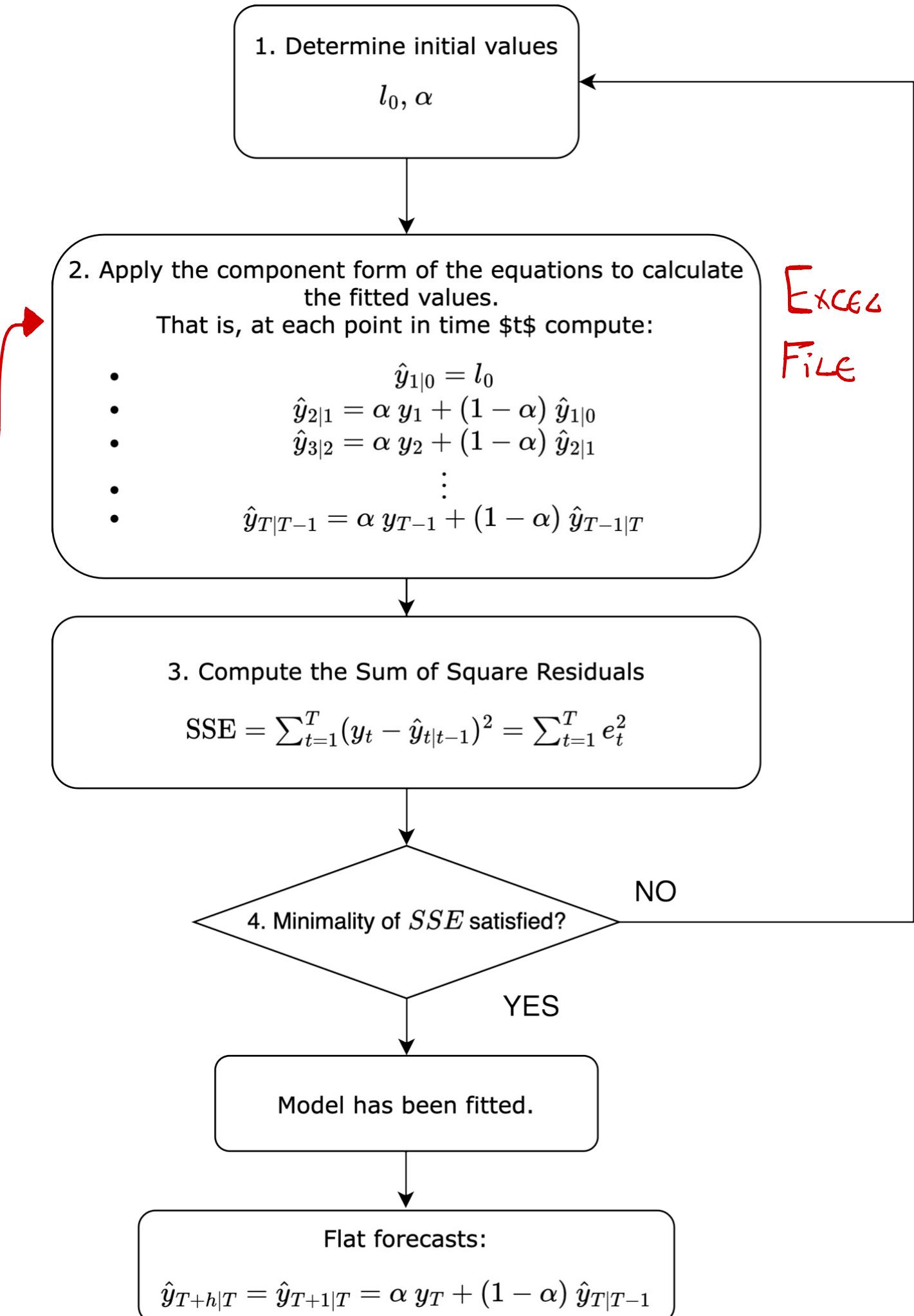
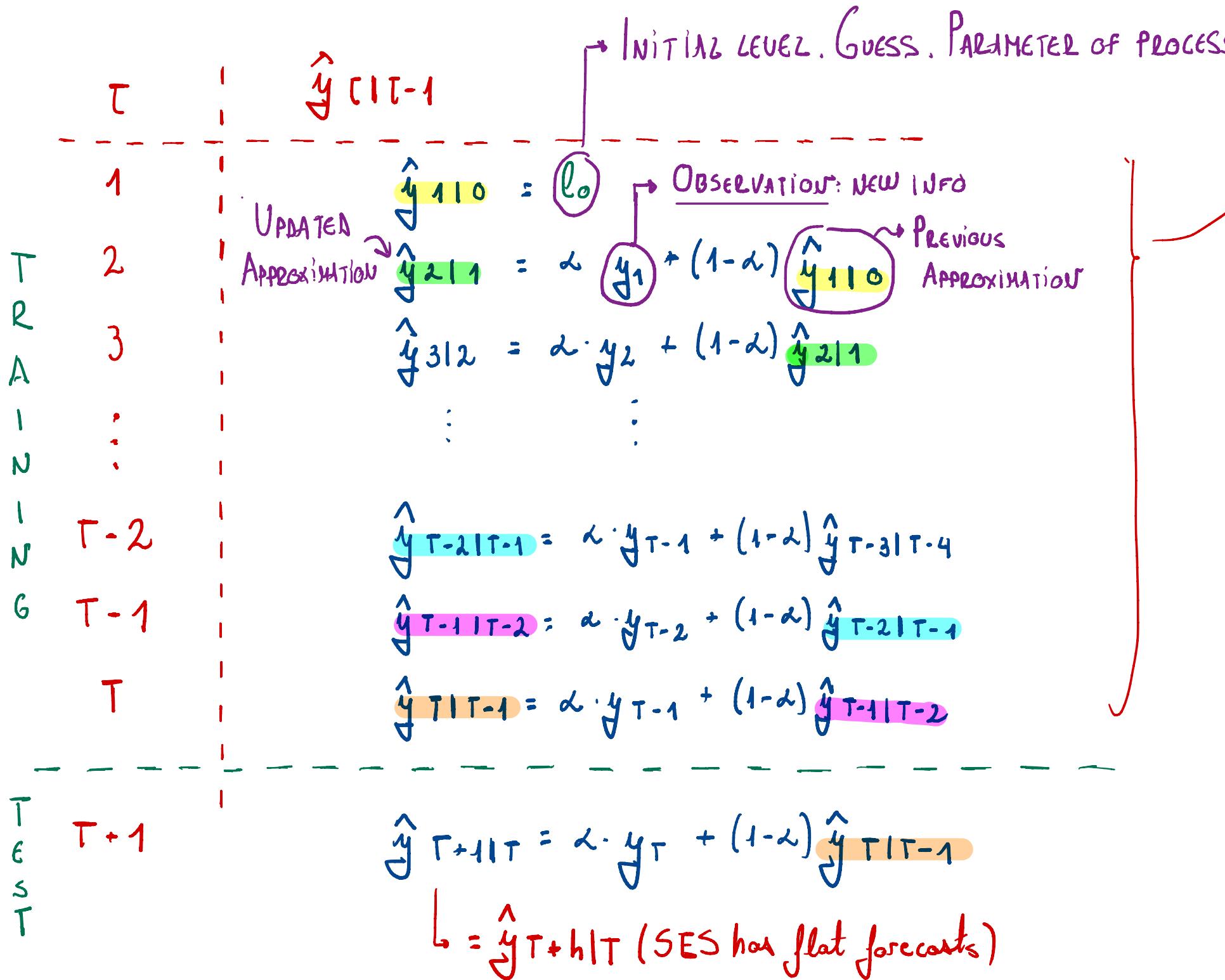
$$\hat{y}_{T+1|T} = \underbrace{\alpha \cdot y_T}_{\text{New APPROX}} + \underbrace{(1-\alpha) \hat{y}_{T|T-1}}_{\substack{\text{New INFO} \\ (\text{OBSERVATION})}}$$

$\hat{y}_{T+1|T}$ PREVIOUS APPROX



FITTING THE MODEL AND GETTING FORECASTS

Given α and l_0 , the equations below provide the fitted values at each point:



PARAMETERS: α AND \underline{l}_0

- SMOOTHING PARAMETER**
- α : determines the weight assigned to the new info (\underline{y}_t) and the previous approximation $((1-\alpha)\hat{y}_{t|t-1})$
 - EQUIVALENT TO: determines how the weight assigned to past observations decays
 - \underline{l}_0 : influences the value of the fitted exponential smoothing
 - Its influence becomes weaker the more distant from $t=1$ the process gets:

$$\begin{array}{l}
 \begin{array}{c|c}
 t & \\
 \hline
 1 & \hat{y}_{1|0} = \underline{l}_0 \\
 2 & \hat{y}_{2|1} = \alpha y_1 + (1-\alpha) \hat{y}_{1|0} = \underbrace{\alpha y_1}_{\text{Weighted by } \alpha} + (1-\alpha) \underline{l}_0 \\
 3 & \hat{y}_{3|2} = \alpha \cdot y_2 + (1-\alpha) \hat{y}_{2|1} = \alpha y_2 + (1-\alpha)[\alpha y_1 + (1-\alpha) \underline{l}_0] = \\
 & = \alpha y_2 + \alpha(1-\alpha) y_1 + \underbrace{(1-\alpha)^2 \underline{l}_0}_{<1} \\
 \vdots & \\
 \vdots & \\
 \text{Generic } t & \hat{y}_{t|t-1} = \alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} + \alpha(1-\alpha)^2 y_{t-3} + \dots + \alpha(1-\alpha)^{t-2} y_0 + \underbrace{(1-\alpha)^{t-1}}_{1-(t-1)} \cdot \underline{l}_0
 \end{array}
 \end{array}$$

WEIGHTED AVERAGE FORM

INFLUENCE OF \underline{l}_0 DECAYS
 EXPONENTIALLY WITH
 INCREASING t

DEFINITION OF THE LEVEL l_T

We define the level at T :
$$l_T = \hat{y}_{T+1|T} \Rightarrow l_0 = \hat{y}_{1|0}$$

$$\begin{array}{c}
 \text{---} \quad | \quad \hat{y}_{T+1|T-1} \\
 T \quad | \quad l_{T-1} \\
 \text{---} \quad | \quad \hat{y}_{T|T-1} = \alpha y_{T-1} + (1-\alpha) \hat{y}_{T-1|T-2} \\
 T-1 \quad | \quad \hat{y}_{T-1|T-2} = \alpha y_{T-2} + (1-\alpha) \hat{y}_{T-2|T-3} \\
 \text{---} \quad | \quad l_{T-2} \\
 T-2 \quad | \quad \vdots \quad \vdots \quad \vdots \quad l_1 \\
 \text{---} \quad | \quad l_2 \\
 3 \quad | \quad \hat{y}_{3|2} = \alpha y_2 + (1-\alpha) \hat{y}_{2|1} \\
 2 \quad | \quad \hat{y}_{2|1} = \alpha y_1 + (1-\alpha) \hat{y}_{1|0} \sim l_0 \\
 1 \quad | \quad \hat{y}_{1|0} = l_0 \quad \text{Initial value for the level}
 \end{array}$$

In summary, with this definition of l_T

$$\begin{array}{c}
 \hat{y}_{T+1|T} = \alpha \cdot y_T + (1-\alpha) \hat{y}_{T|T-1} \\
 l_T = \alpha \cdot y_T + (1-\alpha) l_{T-1}
 \end{array}$$

SMOOTHING EQUATION
 (ALTERNATIVE FORM OF RECURSIVE
 EQUATION FOR FITTED VALUES)

COMPONENT FORM OF THE SES EQUATIONS

FLAT FORECASTS

$$\text{FORECAST EQUATION: } \hat{y}_{T+h|T} = \hat{y}_{T+1|T} = l_T$$

DEFINITION OF LEVEL

$$l_T = \hat{y}_{T+1|T}$$

SMOOTHING EQUATION:

$$\underbrace{\hat{y}_{T+1|T}}_{l_T} = \alpha \cdot y_T + (1-\alpha) \underbrace{\hat{y}_{T|T-1}}_{l_{T-1}} \quad \Leftrightarrow \quad l_T = \alpha \cdot y_T + (1-\alpha) l_{T-1}$$

$$l_T = \hat{y}_{T+1|T} \text{ (DEFINITION OF LEVEL)}$$

SES - EQUATIONS (2 EQUIVALENT FORMS)

WEIGHTED AVERAGE

FORM

$$\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha(1-\alpha)^j y_{T-j}$$

COMPONENT FORM

FORECAST.EQ: $\hat{y}_{T+h|T} = \hat{y}_{T+1|T} = l_T$

SMOOTHING.EQ: $l_T = \alpha \cdot y_T + (1-\alpha) l_{T-1}$

Exponential Moving Average AND ℓ_t

INVESTOPEDIA:

Formula for Exponential Moving Average (EMA)

$$EMA_{\text{Today}} = \left(\text{Value}_{\text{Today}} * \underbrace{\left(\frac{\alpha}{1 + \text{Days}} \right)}_{\text{Smoothing}} \right) + EMA_{\text{Yesterday}} * \underbrace{\left(1 - \left(\frac{\alpha}{1 + \text{Days}} \right) \right)}_{\alpha}$$

e.g.: 20 days $\rightarrow EMA_{20}$

where:

EMA = Exponential moving average

$$EMA_{\text{TODAY}} = \alpha \cdot y_{\text{TODAY}} + (1 - \alpha) EMA_{\text{YESTERDAY}}$$

$$EMA_T = \alpha \cdot y_T + (1 - \alpha) EMA_{T-1}$$

COMPONENT FORM OF SES

$$\text{FORECAST.EQ: } \hat{y}_{T+h|T} = \hat{y}_{T+1|T} = l_T$$

$$\text{SMOOTHING.EQ: } l_T = \alpha \cdot y_T + (1 - \alpha) l_{T-1}$$

The level is the EMA of the time series

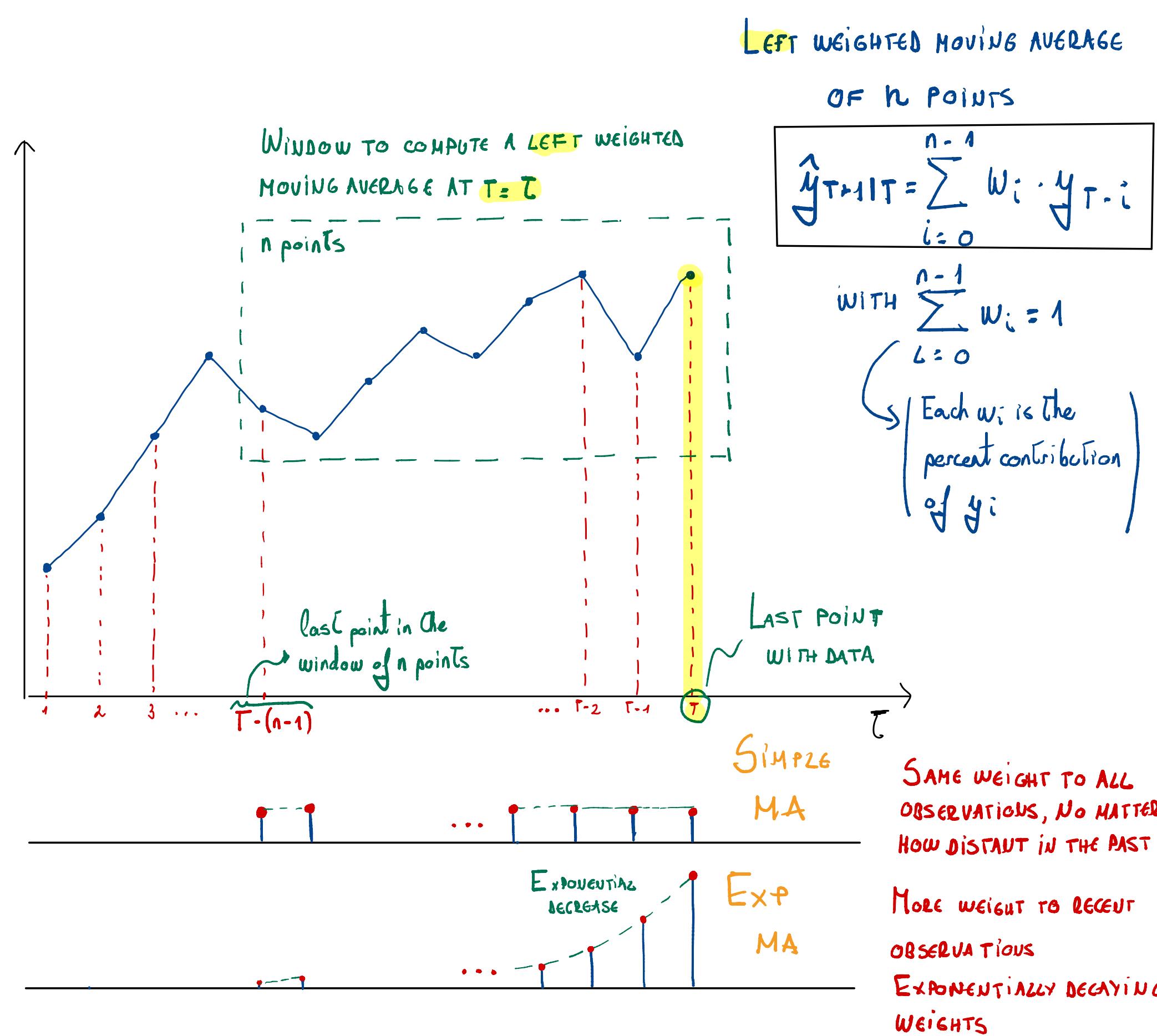
EMA₀ REQUIRED
(l_0)

Calculating the EMA

Calculating the EMA requires one more observation than the SMA. Suppose that you want to use 20 days as the number of observations for the EMA. Then, you must wait until the 20th day to obtain the SMA. On the 21st day, you can then use the SMA from the previous day as the first EMA for yesterday.

SIMPLE MOVING AVERAGE
(NEXT SLIDE)

EXPONENTIAL VS SIMPLE MOVING AVERAGE



- SIMPLE (LEFT) MOVING AVERAGE

$$w_i = \frac{\sum_{j=0}^{n-1} y_j}{n} \quad \forall i \quad \text{→ SAME WEIGHT FOR ALL OBSERVATIONS}$$

- EXPONENTIAL MOVING AVERAGE

$$w_i = \alpha(1-\alpha)^i$$

↳ Obtained from geometric as in SES.

Note: if all points from 1 to T are included in the window ($n = T$)

SMA → Mean Model

EMA → Simple Exponential Smoothing.