Exponential Smoothing

Simon U., Michael Y., Ben H., Sachid Deshmukh

March 3, 2020

Introduction

- **What is exponential smoothing?**
  Forecasting future observations using weighted averages of past observations, with the weights decaying exponentially as observations recede further into the past
  - Presumably, more recent data points are more predictive than older points
  - This framework generates reliable forecasts quickly and for a wide range of time series
- **Exponential Smoothing Models**

1. Naive—all weight is given to the last observation
2. Average—each past observation is given equal weight
3. Exponential weighted average—Recent observations get higher weight, older observations less weight
4. Holt linear—Same as 3, but accounts for time series with trend
5. Holt-Winters—Same as 4, but also accounts for time series with seasonality
6. State Space—ETS models

ES_1: Naive model

- The naive forecasting model can be thought of us exponential smoothing
- Where 100 percent of weight is given to the last observation:

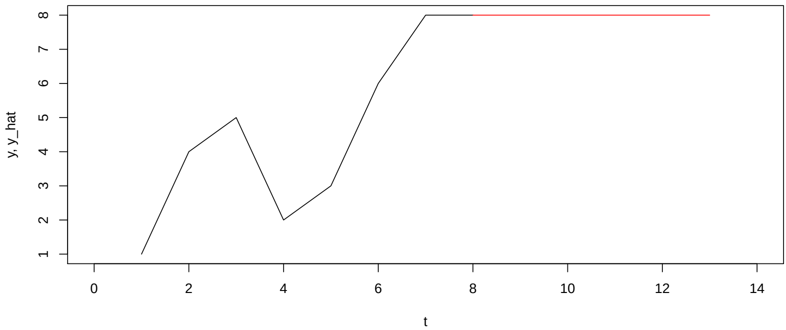
$$\hat{y}_{T+h|T} = y_T$$

```
forecast_naive <- function(y, h) {
    n <- length(y)
    y_hat <- rep(y[n], h)
    return( y_hat )
}
```

ES_1: Naive model: Example

```
y <- c(1, 4, 5, 2, 3, 6, 8)
y_hat <- forecast_naive(y, h=7)

plot_forecast(y, y_hat)
```



ES_2: Average model

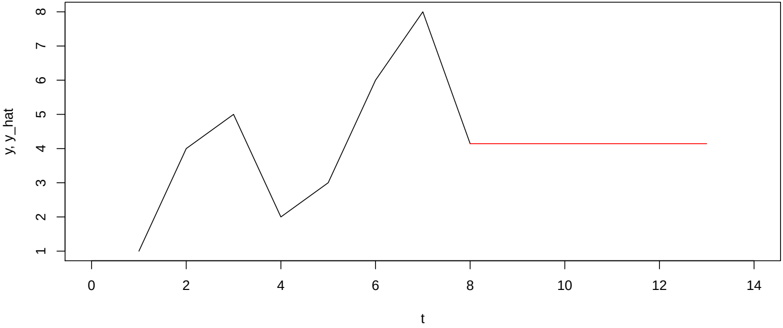
- All future values are forecast as the average of the observed data
- Equivalent to to exponential smoothing where each observation is given equal weight

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} y_t,$$

```
forecast_avg <- function(y, h) {
    y_hat <- rep(mean(y), h)
    return( y_hat )
}
```

ES_2: Average model: Example

```
y_hat <- forecast_avg(y, h=7)
plot_forecast(y, y_hat)
```



ES_3: Simple Exponential Smoothing [SES]

- SES stands at the core and servers as the foundation for some of the most successful forecasting methods for modeling time series data. The produced forecasts are weighted averages of past observations, with the weights decaying exponentially as the observations get older. SES is actually suitable for forecasting data with no clear trend or seasonal pattern.

Exploring "Exponential" aspect

- The formula for estimating a value at $T+1$ is as follows:

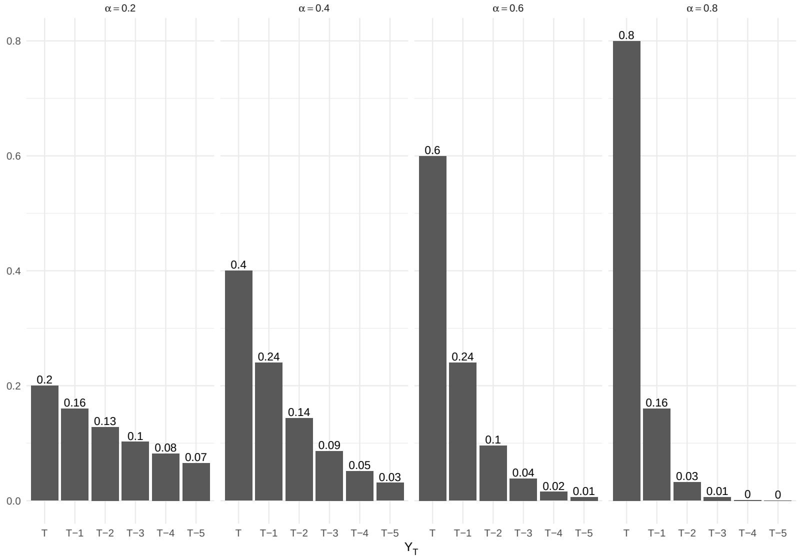
$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots$$

Geometric Distribution

- If the probability of a success in one trial is p and the probability of a failure is $1-p$, then the probability of finding the first success in the n^{th} tiral is given by

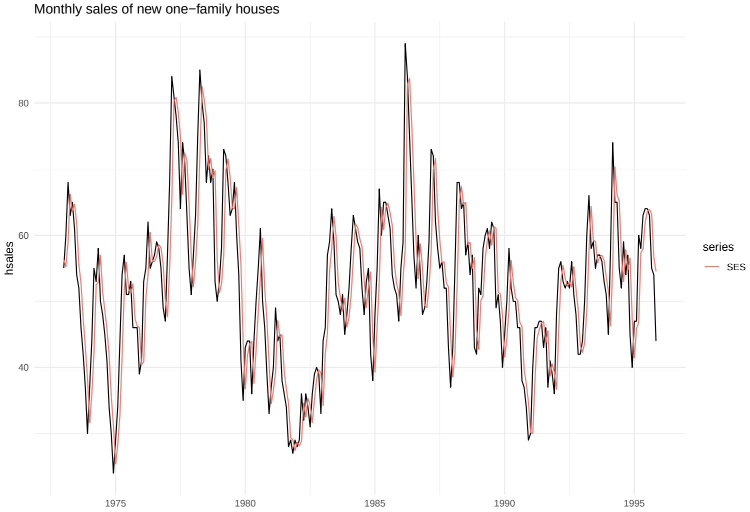
$$p(1-p)^{n-1}$$

Geometric Distribution graph for various values of α

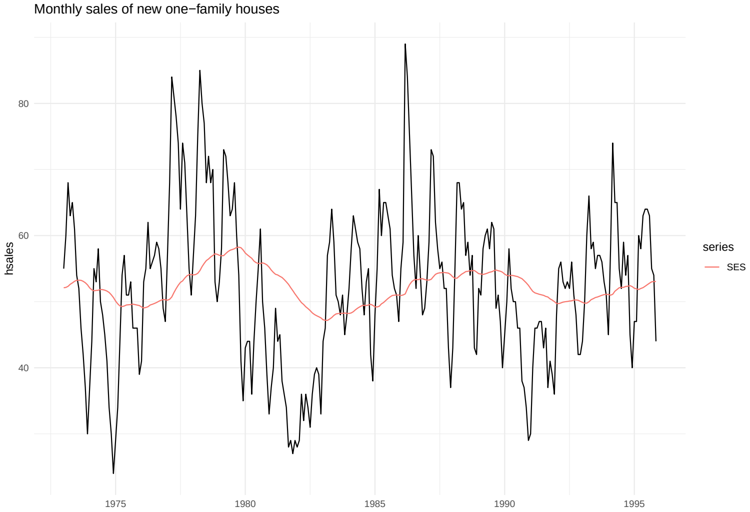


Exploring "Smoothing" aspect

- With $\alpha \approx 1$ [$\alpha = 0.8$],
  the model is very much ***reactive*** to the most recent observations



- With $\alpha \approx 0$ [$\alpha = 0.02$], the model is much less reactive (***smoother***) to react to change



SES - Mathematical Formulations

- **Weighted average form**

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

This is a *"recursive"* definition for the below

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

which can also be written as:

$$\hat{y}_{T+1|T} = \sum_{j=0}^{T-1} \alpha(1-\alpha)^j y_{T-j} + (1-\alpha)^T \ell_0$$

- **Component Form**

$$\begin{aligned} \text{Forecast equation} \quad & \hat{y}_{t+h|t} = \ell_t \\ \text{Smoothing equation} \quad & \ell_t = \alpha y_t + (1-\alpha)\ell_{t-1} \end{aligned}$$

where ℓ_t is the ***level*** (*smoothed* value) of the series at time t

Interlude

- The previous method is effective for time series without trend or seasonality
- But what if your time series has trend?

ES_4: Holt Linear Trend Model

Appropriate for time series that can be described with

$$y_t = \beta_0 + \beta_1 t + \epsilon_t$$

where β_1 quantifies the trend

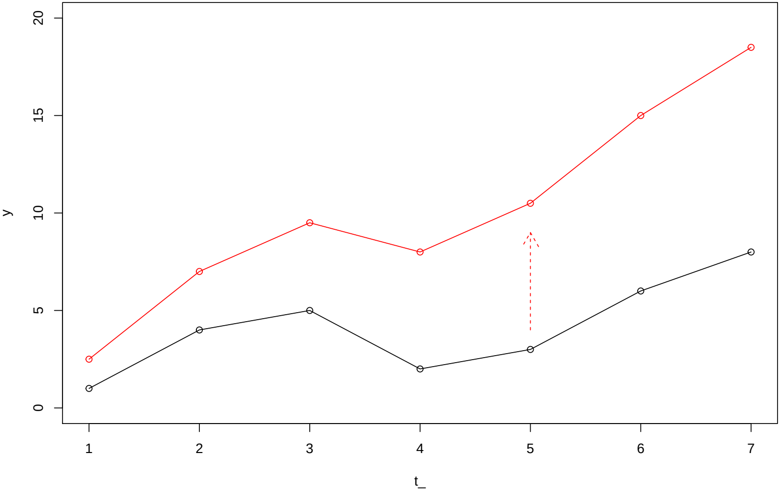
ES_4: Holt Linear Trend Model: Linear Time Series

- We can explicitly convert our previous time series to have trend using this formula:

```
beta_0 <- 0
beta_1 <- 1.5
t_ <- 1:7
( y_1 <- beta_0 + y + beta_1*t_ )
```

```
## [1]  2.5  7.0  9.5  8.0 10.5 15.0 18.5
```

ES_4: Holt Linear Trend Model: Linear Time Series



ES_4: Holt Linear Trend Model: Function/equations

TODO

ES_4: Holt Linear Trend Model: Linear Time Series: Problem

Problem: What if you expect the trend to 'stabilize' over time?

Modify the above equations to use a *dampening* coefficient $0 < \phi < 1$:

Holt-Winters method + multiplicative + taxonomy

ETS modeling (Innovations state space models)

Conclusion

Resources & Further Readings

- Exponential Smoothing
  - https://otexts.com/fpp2/expsmooth.html
  - http://uc-r.github.io/ts_exp_smoothing
- Geometric Distribution
  - OpenIntro Statistics, Fourth Edition, David Diez, Mine Cetinkaya-Rundel, Christopher D Barr