Exponential Smoothing

Simon U., Michael Y., Ben H.

March X, 2020

What is exponential smoothing?

- Forecasting future observations
- Using weighted averages of past observations
  - weights decay exponentially as observations recede further into the past
- Basic idea: More recent observations are more predictive than older points

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—

ES_1: Naive model

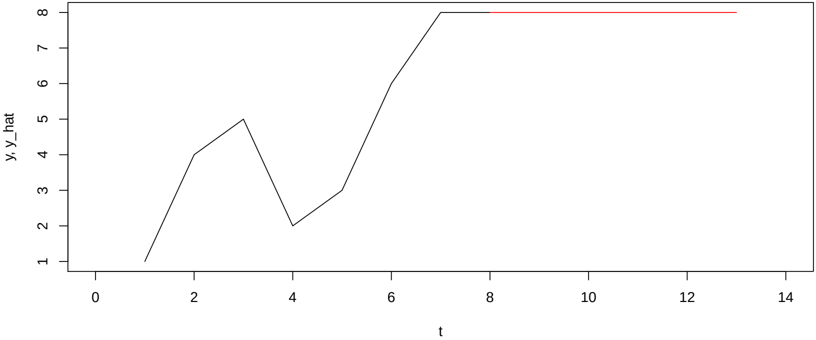
- The naive forecasting model can be thought of us exponential smoothing
- Where 100 percent of weight is given to the last observation:

```
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)
```



Note: No trend or seasonality!

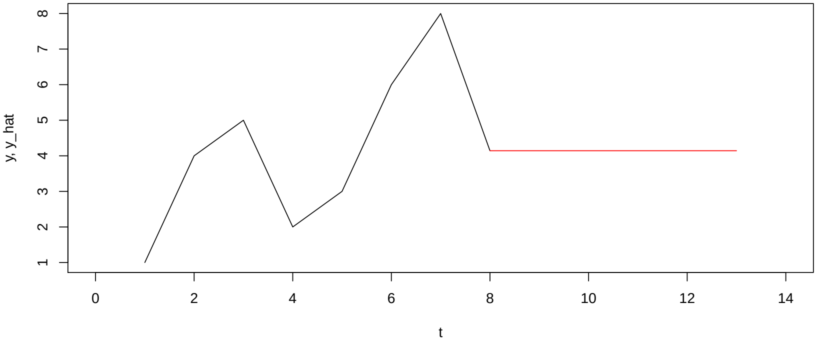
ES_2: Average model

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

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

ES_2: Average model: Example

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



ES_3: Simple exponential smoothing

- More sophisticated models would given recent observations more weight, and decreasing weight for past observations
- Parameter α controls smoothing; can be optimized

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

- $\hat{y}_T \equiv$ predicted value of y at time t
- $\alpha \equiv$ user-chosen smoothing parameter, $0 \leq \alpha \leq 1$
  - Closer to 0 gives historical data more weight, closer to 1 gives recent data more weight
  - Often between $0.1 - -0.2$ is best
- $\hat{y}_{T-1} \equiv$ predicted value of y at immediately previous period $t - 1$

ES_3: Simple exponential smoothing: Smoothing parameter

where y_t is the most recent observation:

Observation	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
y_t	0.2	0.4	0.6	0.8
y_{t-1}	0.16	0.24	0.26	0.16
y_{t-2}	0.128	0.144	0.096	0.032
y_{t-3}	0.1024	0.0864	0.0384	0.0064
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

ES_3: Simple exponential smoothing: Problem

Problem: What is $\hat{y}_{T-1}$ when $t = 1$?

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

- Textbook: Component form, the level ℓ_0
- Set for mean or median of the time series

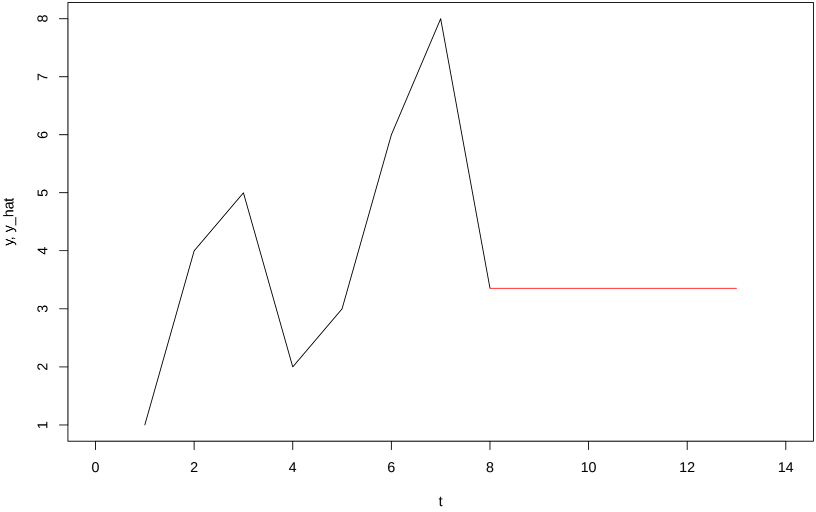
```r
forecast_simple <- function(y, h, alpha=0.1) {
    n <- length(y)
    l_0 <- mean(y)

    l <- c()
    for (i in 1:(n+1)) {
        if (i == 1) {
            l_i <- alpha * y[i] + (1 - alpha) * l_0
        } else {
            l_i <- alpha * y[i] + (1 - alpha) * l[i-1]
        }
        l <- append(l_i, l)
    }

    y_hat <- l[n+1]
    return( rep(y_hat, h) )
}
```

ES_3: Simple exponential smoothing: Example

```
y_hat <- forecast_simple(y, h=7, alpha=0.25)
plot_forecast(y, y_hat)
```



ES_3: Weighted average: Optimizing α^*

BEN: Not sure if we want to keep this??

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

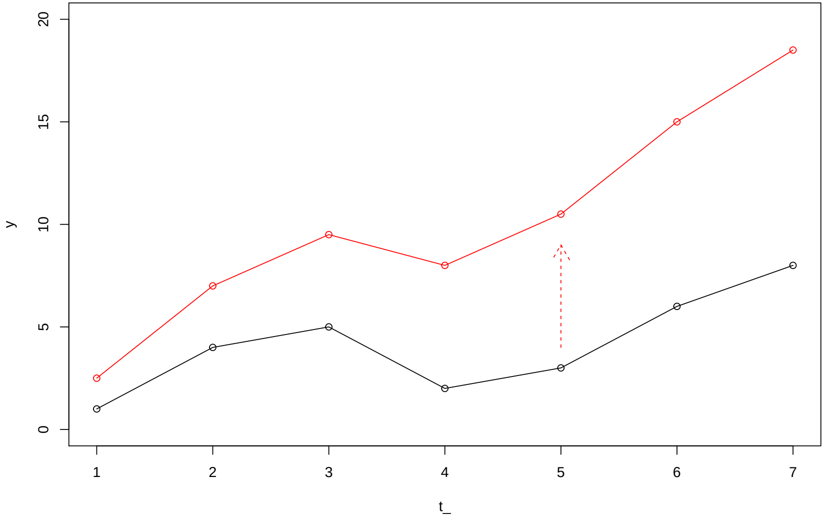
ES_4: Holt Linear Trend Model: Linear Time Series

- 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

Blah blah