MONASH University

MONASH BUSINESS SCHOOL

ETF3231/5231 Business forecasting

Week 5: Exponential smoothing

`https://bf.numbat.space/`

AACSB ACCREDITED

EQUIS ACCREDITED

ASSOCIATION OF MBAS AMBA ACCREDITED

Outline

Outline

Historical perspective

Developed in the US navy for forecasting spare parts

- Proposed in the late 1950s (Brown 1959, Holt 1957 and Winters 1960 are key pioneering works) as methods (algorithms) to produce point forecasts.

  for this reason not popular with statisticians.

KEY IDEA

- Forecasts are weighted averages of past observations, with the weights decaying smoothly exponentially as the observations get older.
- Framework generates reliable forecasts quickly and for a wide spectrum of time series. A great advantage and of major importance to applications in industry.

  - now used everywhere in business
  - strong benchmarks

Combine components

- Combine components: level ℓ_t , trend (slope) b_t and seasonal s_t to describe a time series

$$y_t = f(\ell_{\cancel{t}\,t-1}, b_{\cancel{t}\,t-1}, s_{\cancel{t}\,t-m}) \longrightarrow \hat{y}_{T+h|T} = f(\ell_T, b_T, s_{T-m+1})$$

- The rate of change of the components are controlled by "smoothing parameters": α, β and γ respectively. see next slide
- Need to choose best values for the smoothing parameters (and initial states).
- Add error ε_t to get equivalent ETS state space models developed in the 1990s and 2000s.

* Monash EBS very famous about these
* Pioneer Ralph Snyder (textbook with Rob Hyndman Ann Koehler & Keith Ord).

Big idea: control the rate of change (smoothing)

(height, overall position of the series)

α controls the flexibility of the level ℓ_t

- If α = 0, the level never updates (mean)
- If α = 1, the level updates completely (naive)

(slope)

β controls the flexibility of the trend b_t

- If β = 0, the trend is linear (regression trend)
- If β = 1, the trend updates every observation

usually

$0 \le \alpha, \beta, \gamma \le 1$

(more to follow)

γ controls the flexibility of the seasonality s_t

- If γ = 0, the seasonality is fixed (seasonal means)
- If γ = 1, the seasonality updates completely (seasonal naive)

A model for levels, trends, and seasonalities

We want a model that captures the level (ℓ_t), trend (b_t) and seasonality (s_t).

How do we combine these elements?

A model for levels, trends, and seasonalities

We want a model that captures the level (ℓ_t), trend (b_t) and seasonality (s_t).

How do we combine these elements?

Additively?

$$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$$

$\varepsilon_t \sim \text{iid } N(0, \sigma^2)$

Multiplicatively?

$$y_t = \ell_{t-1} b_{t-1} s_{t-m} (1 + \varepsilon_t)$$

Perhaps a mix of both?

$$y_t = (\ell_{t-1} + b_{t-1}) s_{t-m} + \varepsilon_t$$

A model for levels, trends, and seasonalities

We want a model that captures the level (ℓ_t), trend (b_t) and seasonality (s_t).

How do we combine these elements?

Additively?

$$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$$

Multiplicatively?

$$y_t = \ell_{t-1} b_{t-1} s_{t-m}(1 + \varepsilon_t)$$

Perhaps a mix of both?

$$y_t = (\ell_{t-1} + b_{t-1}) s_{t-m} + \varepsilon_t$$

How do the level, trend and seasonal components evolve over time?

* Hence more than one equation gets used.

ETS models

General notation E T S : **E**xponen**T**ial **S**moothing

Error **T**rend **S**eason

```
model(ETS(y ~ error() + trend() + season()))
```

ETS models

General notation E T S : **E**xponen**T**ial **S**moothing

Error Trend Season

```
model(ETS(y ~ error() + trend() + season()))
```

Error: Additive (`"A"`) or multiplicative (`"M"`)

ETS models

General notation E T S : **E**xponen**T**ial **S**moothing

Error Trend Season

```
model(ETS(y ~ error() + trend() + season()))
```

Error: Additive (`"A"`) or multiplicative (`"M"`)

Trend: None (`"N"`), additive (`"A"`), multiplicative (`"M"`), or damped (`"Ad"` or `"Md"`).

ETS models

General notation E T S : **E**xponen**T**ial **S**moothing

Error Trend Season

```
model(ETS(y ~ error() + trend() + season()))
```

Error: Additive (`"A"`) or multiplicative (`"M"`)

Trend: None (`"N"`), additive (`"A"`), multiplicative (`"M"`), or damped (`"Ad"` or `"Md"`).

Seasonality: None (`"N"`), additive (`"A"`) or multiplicative (`"M"`)

- Hence many combinations of these (in theory 30 models, in practice about half of these)

Models and methods

Methods

- Algorithms that return point forecasts.

Models and methods

Methods

- Algorithms that return point forecasts.

Models

- Generate same point forecasts but can also generate forecast distributions.
- A stochastic (or random) data generating process that can generate an entire forecast distribution.

  Can generate something that looks like data
- Allow for "proper" model selection.

· Ord, Koehler, Snyder (1997, JASA)

* machine learning methods (NNs) are in their algorithmic phase.

Outline

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha)\hat{y}_{1|0}$$

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha)\cancel{\hat{y}_{1|0}}\ \ell_0$$

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha) \cancel{\hat{y}_{1|0}} \; l_0$$

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

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha)\cancel{\hat{y}_{1|0}}\ \ell_0$$

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

$$\hat{y}_{4|3} = \alpha y_3 + (1-\alpha)\hat{y}_{3|2}$$

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1-\alpha)\hat{y}_{t|t-1}$$

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha)\cancel{\hat{y}_{1|0}}\ \ell_0$$

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

$$\hat{y}_{4|3} = \alpha y_3 + (1-\alpha)\hat{y}_{3|2}$$

$$\vdots$$

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

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

Weighted average form

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

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

$$= \alpha y_T + (1-\alpha)\left[\alpha y_{T-1} + (1-\alpha)\hat{y}_{T-1|T-2}\right]$$

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

$$= \alpha y_T + \alpha(1-\alpha) y_{T-1} + (1-\alpha)^2\left[\alpha y_{T-2} + (1-\alpha)\hat{y}_{T-2|T-3}\right]$$

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

$$\vdots$$

$$= \alpha y_T + \alpha(1-\alpha) y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots + (1-\alpha)^T \ell_0$$

we don't have infinite data.

when $\alpha = 1$ $\quad \hat{y}_{T+1|T} = y_T$ $\quad \rightarrow$ only last obs matters

$\alpha = 0$ $\quad \hat{y}_{T+1|T} = \ell_0$ $\quad \rightarrow$ we learn nothing from new info

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

Weighted average form

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

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

Weighted average form

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

Component form

Forecast equation	$\hat{y}_{t+1\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

Simple Exponential Smoothing - SES

Iterative form

$$\hat{y}_{t+1|t} = \alpha y_t + (1-\alpha)\hat{y}_{t|t-1}$$

let $\hat{y}_{t+1|t} = \ell_t$

$\Rightarrow \hat{y}_{t|t-1} = \ell_{t-1}$

Weighted average form

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

Component form

Holt

Forecast equation $\quad \hat{y}_{t+1|t} = \ell_t$

Smoothing equation $\quad \ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

ETS(A,N,N): SES with additive errors

Component form

Forecast equation	$\hat{y}_{t+1\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

ETS(A,N,N): SES with additive errors

Component form

Forecast equation	$\hat{y}_{t+1\|t} = \ell_t$
Smoothing equation	$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

Residual: $e_t = y_t - \hat{y}_{t|t-1} = y_t - \ell_{t-1}$.

ETS(A,N,N): SES with additive errors

Component form

Forecast equation $\hat{y}_{t+1|t} = \ell_t$

Smoothing equation $\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

Residual: $e_t = y_t - \hat{y}_{t|t-1} = y_t - \ell_{t-1}$.

Error correction form

$$\begin{aligned} y_t &= \ell_{t-1} + e_t \\ \ell_t &= \ell_{t-1} + \alpha(y_t - \ell_{t-1}) \\ &= \ell_{t-1} + \alpha e_t \end{aligned}$$

ETS(A,N,N): SES with additive errors

Component form

Forecast equation $\hat{y}_{t+1|t} = \ell_t$

Smoothing equation $\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1}$

Residual: $e_t = y_t - \hat{y}_{t|t-1} = y_t - \ell_{t-1}$.

Error correction form

$$y_t = \ell_{t-1} + e_t$$

$$\begin{aligned} \ell_t &= \ell_{t-1} + \alpha(y_t - \ell_{t-1}) \\ &= \ell_{t-1} + \alpha e_t \end{aligned}$$

KEY RESULT

Specify probability distribution for e_t, we assume $e_t = \varepsilon_t \sim \text{NID}(0, \sigma^2)$.

ETS(A,N,N): SES with additive errors

An ETS model

Measurement equation	$y_t = \ell_{t-1} + \varepsilon_t$
State equation	$\ell_t = \ell_{t-1} + \alpha\varepsilon_t$

where $\varepsilon_t \sim \text{NID}(0, \sigma^2)$.

- innovations or single source of error because equations have the same error process, ε_t.
- Measurement or observation ~~equation~~: relationship between observations and states.
- State equation(s): evolution of the state(s) through time.

ETS(A,N,N): SES with additive errors

An ETS model

Measurement equation	$y_t = \ell_{t-1} + \varepsilon_t$
State equation	$\ell_t = \ell_{t-1} + \alpha\varepsilon_t$

where $\varepsilon_t \sim \text{NID}(0, \sigma^2)$.

- innovations or single source of error because equations have the same error process, ε_t.
- Measurement ~~equation~~ or observation equation: relationship between observations and states.
- State equation(s): evolution of the state(s) through time.

QUESTION / HOMEWORK : what happen when $\alpha = 1$?

ETS(M,N,N): SES with multiplicative errors.

- Specify relative errors $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- Substituting $\hat{y}_{t|t-1} = \ell_{t-1}$ gives:
  - $y_t = \ell_{t-1} + \ell_{t-1}\varepsilon_t$
  - $e_t = y_t - \hat{y}_{t|t-1} = \ell_{t-1}\varepsilon_t$

ETS(M,N,N): SES with multiplicative errors.

- Specify relative errors $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- Substituting $\hat{y}_{t|t-1} = \ell_{t-1}$ gives:
  - $y_t = \ell_{t-1} + \ell_{t-1}\varepsilon_t$
  - $e_t = y_t - \hat{y}_{t|t-1} = \ell_{t-1}\varepsilon_t$

Measurement equation	$y_t = \ell_{t-1}(1 + \varepsilon_t)$
State equation	$\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$

ETS(M,N,N): SES with multiplicative errors.

- Specify relative errors $\varepsilon_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}} \sim \text{NID}(0, \sigma^2)$
- Substituting $\hat{y}_{t|t-1} = \ell_{t-1}$ gives:
  - $y_t = \ell_{t-1} + \ell_{t-1}\varepsilon_t$
  - $e_t = y_t - \hat{y}_{t|t-1} = \ell_{t-1}\varepsilon_t$

Measurement equation	$y_t = \ell_{t-1}(1 + \varepsilon_t)$
State equation	$\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$

they are different models

- Models with additive and multiplicative errors with the same parameters generate the same point forecasts but different prediction intervals.

(ℓ_0, α)