# ETC3550/ETC5550
# Applied forecasting

Ch8. Exponential smoothing

OTexts.org/fpp3/

# ETS models

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

**E**rror  **T**rend  **S**eason

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

# ETS models

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

**E**rror  **T**rend  **S**eason

**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

**E**rror   **T**rend   **S**eason

**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"`)

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

## ETS(A,N,N) model

| | |
|---|---|
| Observation 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, $\varepsilon_t$.
- Measurement equation: relationship between observations and states.
- State equation(s): evolution of the state(s) through time.

# ETS(A,A,N)

Holt's methods method with additive errors.

| | |
|---|---|
| Forecast equation | $\hat{y}_{t+h\|t} = \ell_t + hb_t$ |
| Observation equation | $y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ |
| State equations | $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ |
| | $b_t = b_{t-1} + \beta\varepsilon_t$ |

- Forecast errors: $\varepsilon_t = y_t - \hat{y}_{t\|t-1}$

# ETS(A,A,A)

Holt-Winters additive method with additive errors.

| | |
|---|---|
| Forecast equation | $\hat{y}_{t+h|t} = \ell_t + h b_t + s_{t+h-m(k+1)}$ |
| Observation equation | $y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ |
| State equations | $\ell_t = \ell_{t-1} + b_{t-1} + \alpha \varepsilon_t$ |
| | $b_t = b_{t-1} + \beta \varepsilon_t$ |
| | $s_t = s_{t-m} + \gamma \varepsilon_t$ |

- Forecast errors: $\varepsilon_t = y_t - \hat{y}_{t|t-1}$
- $k$ is integer part of $(h-1)/m$.

# ETS(M,A,M)

Holt-Winters multiplicative method with multiplicative errors.

| | |
|---|---|
| Forecast equation | $\hat{y}_{t+h\|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$ |
| Observation equation | $y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$ |
| State equations | $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ |
| | $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ |
| | $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$ |

- Forecast errors: $\varepsilon_t = (y_t - \hat{y}_{t\|t-1})/\hat{y}_{t\|t-1}$
- $k$ is integer part of $(h - 1)/m$.

## ETS model specification

```
ETS(y ~ error("A") + trend("N") + season("N"))
```

By default, optimal values for $\alpha$, $\beta$, $\gamma$, and the states at time 0 are used.

The values for $\alpha$, $\beta$ and $\gamma$ can be specified:

```
trend("A", alpha = 0.5, beta = 0.2)
trend("A", alpha_range = c(0.2, 0.8), beta_range = c(0.1, 0.4))
season("M", gamma = 0.04)
season("M", gamma_range = c(0, 0.3))
```

# Exponential smoothing methods

| Trend Component | Seasonal Component | | |
|---|---|---|---|
| | N (None) | A (Additive) | M (Multiplicative) |
| N (None) | (N,N) | (N,A) | (N,M) |
| A (Additive) | (A,N) | (A,A) | (A,M) |
| $A_d$ (Additive damped) | $(A_d,N)$ | $(A_d,A)$ | $(A_d,M)$ |

**(N,N)**: Simple exponential smoothing
**(A,N)**: Holt's linear method
**$(A_d,N)$**: Additive damped trend method
**(A,A)**: Additive Holt-Winters' method
**(A,M)**: Multiplicative Holt-Winters' method
**$(A_d,M)$**: Damped multiplicative Holt-Winters' method

# Exponential smoothing methods

|  | | Seasonal Component | | |
|---|---|---|---|---|
| Trend Component | | N (None) | A (Additive) | M (Multiplicative) |
| N | (None) | (N,N) | (N,A) | (N,M) |
| A | (Additive) | (A,N) | (A,A) | (A,M) |
| $A_d$ | (Additive damped) | $(A_d,N)$ | $(A_d,A)$ | $(A_d,M)$ |

**(N,N)**:  Simple exponential smoothing
**(A,N)**:  Holt's linear method
**($A_d$,N)**:  Additive damped trend method
**(A,A)**:  Additive Holt-Winters' method
**(A,M)**:  Multiplicative Holt-Winters' method
**($A_d$,M)**:  Damped multiplicative Holt-Winters' method

There are also multiplicative trend methods (not recommended).

# ETS models

| Additive Error | | Seasonal Component | | |
|---|---|---|---|---|
| **Trend** | | N | A | M |
| **Component** | | (None) | (Additive) | (Multiplicative) |
| N | (None) | A,N,N | A,N,A | A,N,M |
| A | (Additive) | A,A,N | A,A,A | A,A,M |
| $A_d$ | (Additive damped) | A,$A_d$,N | A,$A_d$,A | A,$A_d$,M |

| Multiplicative Error | | Seasonal Component | | |
|---|---|---|---|---|
| **Trend** | | N | A | M |
| **Component** | | (None) | (Additive) | (Multiplicative) |
| N | (None) | M,N,N | M,N,A | M,N,M |
| A | (Additive) | M,A,N | M,A,A | M,A,M |
| $A_d$ | (Additive damped) | M,$A_d$,N | M,$A_d$,A | M,$A_d$,M |

# ETS models

| **Additive Error** | | **Seasonal Component** | | |
|---|---|---|---|---|
| **Trend** | | N | A | M |
| **Component** | | (None) | (Additive) | (Multiplicative) |
| N | (None) | A,N,N | A,N,A | ~~A,N,M~~ |
| A | (Additive) | A,A,N | A,A,A | ~~A,A,M~~ |
| $A_d$ | (Additive damped) | A,$A_d$,N | A,$A_d$,A | ~~A,$A_d$,M~~ |

| **Multiplicative Error** | | **Seasonal Component** | | |
|---|---|---|---|---|
| **Trend** | | N | A | M |
| **Component** | | (None) | (Additive) | (Multiplicative) |
| N | (None) | M,N,N | M,N,A | M,N,M |
| A | (Additive) | M,A,N | M,A,A | M,A,M |
| $A_d$ | (Additive damped) | M,$A_d$,N | M,$A_d$,A | M,$A_d$,M |

Minimizing the AIC assuming Gaussian residuals is asymptotically equivalent to minimizing one-step time series cross validation MSE.

# Automatic forecasting

**From Hyndman et al. (IJF, 2002):**

- Apply each model that is appropriate to the data. Optimize parameters and initial values using MLE (or some other criterion).
- Select best method using AICc:
- Produce forecasts using best method.
- Obtain forecast intervals using underlying state space model.

Method performed very well in M3 competition.

# Residuals

## Response residuals

$$\hat{e}_t = y_t - \hat{y}_{t|t-1}$$

## Innovation residuals

Additive error model:

$$\hat{\varepsilon}_t = y_t - \hat{y}_{t|t-1}$$

Multiplicative error model:

$$\hat{\varepsilon}_t = \frac{y_t - \hat{y}_{t|t-1}}{\hat{y}_{t|t-1}}$$