# Week 4: Holt's method, seasonality, periodicity

# Holt's method

Situation: $Y_t = Tr_t + E_t$

Assumption: Locally linear trend

$$Tr_{t+j} = L_t + T_t \cdot j$$

$L_t$ level, $T_t$ slope

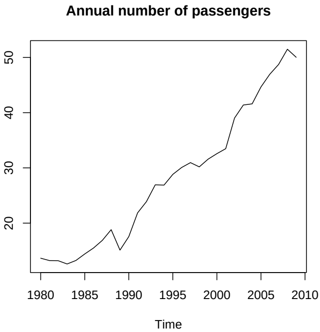
# Holt's method

Situation: $Y_t = Tr_t + E_t$

Assumption: Locally linear trend

$$Tr_{t+j} = L_t + T_t \cdot j$$

$L_t$ level, $T_t$ slope



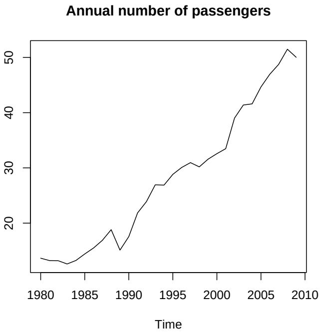
# Holt's method

Situation: $Y_t = Tr_t + E_t$

Assumption: Locally linear trend

$$Tr_{t+j} = L_t + T_t \cdot j$$

$L_t$ level, $T_t$ slope

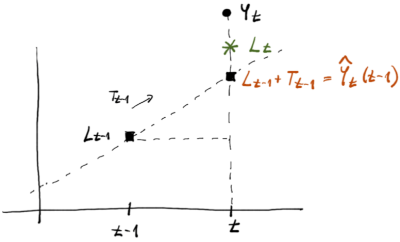


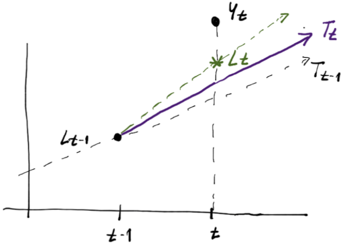

- two smoothing parameters $\alpha$, $\gamma$
- sequentially computes $L_t$, $T_t$ using recursive formulas

# Holt's method

$$\text{Level equation} \quad L_t = \alpha Y_t + (1-\alpha)\underbrace{(L_{t-1} + T_{t-1})}_{\widehat{Y}_t(t-1)},$$

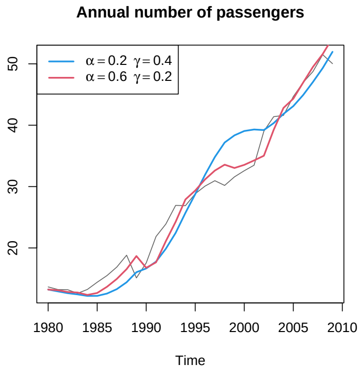
$$\text{Trend equation} \quad T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$$



# Parameters

- Choice of $\alpha, \gamma$:
  - minimizing the SSE
- Choice of $L_0, T_0$:
  - $L_0 = Y_1,\ T_0 = Y_2 - Y_1$,
  - regression on $K$ first observations
  - minimizing SSE

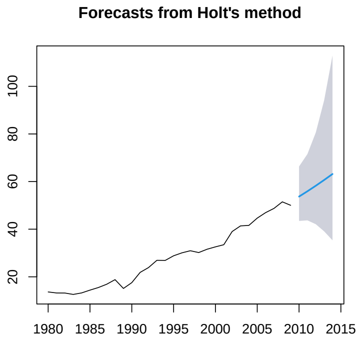


# Forecast

- Forecast

$$\widehat{Y}_{n+h}(t) = L_n + T_n \cdot h$$

- additional assumptions $\rightsquigarrow$ prediction interval



# Double exponential smoothing and Holt's method

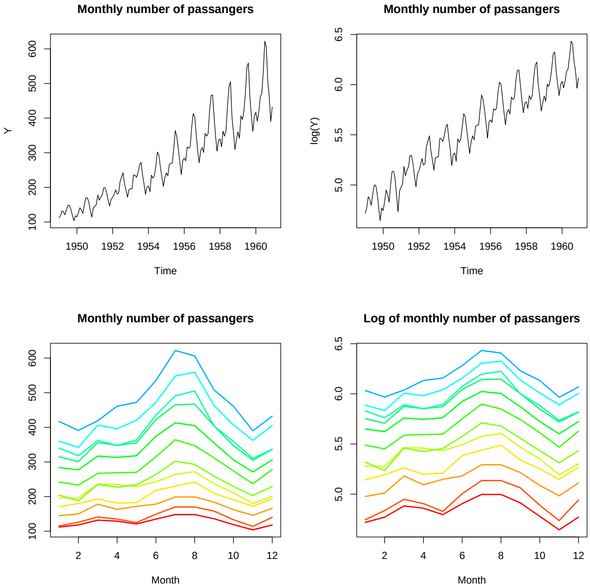
Double exp. smoothing with parameter $a \leftrightsquigarrow$ Holt's method for

$$\alpha_{Holt} = a(2-a), \quad \gamma_{Holt} = \frac{a}{2-a},$$

# Modelling seasonality

# Seasonality

$\{Y_t\}$ exhibits regular periodic changes over a year



# Models

- additive

$$Y_t = Tr_t + S_t + E_t$$

- multiplicative

$$Y_t = Tr_t \cdot S_t \cdot E_t$$

  (use log to get an additive model)

Let $s$ be the number of seasons within a time unit

- $s = 12$ for monthly data, $s = 4$ for quarterly data (time unit =year)
- ...

# Seasonal indices

Seasonal indices: $I_1, \ldots, I_s \rightsquigarrow$ describe seasonal pattern

**Additive model** $Y_t = Tr_t + S_t + E_t$:

- $$I_1 + \cdots + I_s = 0$$

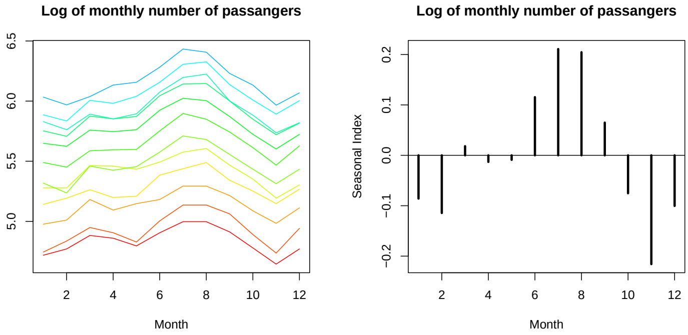
  $I_i$: how season $i$ differs from the overall yearly average
- $I_i$ measured in the same units as $Y_t$

**Multiplicative model** $Y_t = Tr_t \cdot S_t \cdot E_t$

- $$I_1 \cdot \cdots \cdot I_s = 1$$

  (see what happens when taking log)
- $I_i$ relative variable

# Example


Log of monthly number of passangers
6.5
6.0
5.5
5.0
2
4
6
8
10
12
Month
Log of monthly number of passangers
Seasonal Index
0.2
0.1
0.0
−0.1
−0.2
2
4
6
8
10
12
Month

# Approaches to seasonality

1. Simple decomposition
2. Regression approach
3. Smoothing: Holt-Winter's method

# Simple decomposition

$$Y_t = Tr_t + S_t + E_t$$

For monthly data ($s = 12$):

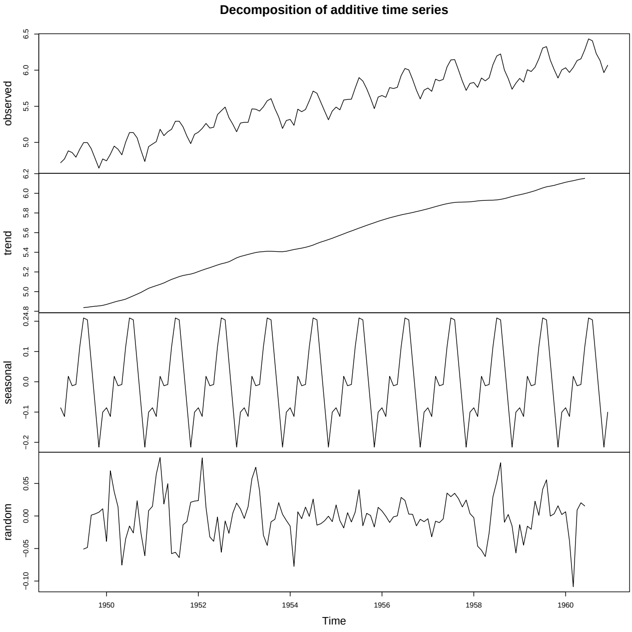
- compute centered moving averages ($m = 6$) $\rightsquigarrow \widehat{Tr}_t$
- take

  $$Y_t - \widehat{Tr}_t$$

  and compute average for each month $\rightsquigarrow \widehat{S}_1, \ldots, \widehat{S}_{12}$
- compute seasonal indices

  $$I_j = \widehat{S}_j - \frac{1}{12} \sum_{i=1}^{12} \widehat{S}_j$$

- if required, seasonally adjusted series: $Y_t - I_j$, where $j$ corresponds to the month of $t$


Decomposition of additive time series
observed
trend
seasonal
random
Time
1950
1952
1954
1956
1958
1960

# Regression approaches

$$Y_t = \beta_0 + \beta_1 t + S_t + E_t$$

Model for $S_t$ (for monthly data $s = 12$):

- using **dummy variables**
  - factor with 12 levels $\rightsquigarrow \widehat{S}_j$, $j = 1, \ldots, 12$
  - recompute to $I_j$

# Regression approaches

$$Y_t = \beta_0 + \beta_1 t + S_t + E_t$$

Model for $S_t$ (for monthly data $s = 12$):

- using **dummy variables**
  - factor with 12 levels $\rightsquigarrow \widehat{S}_j$, $j = 1, \dots, 12$
  - recompute to $I_j$
- using **goniometric functions**, e.g.

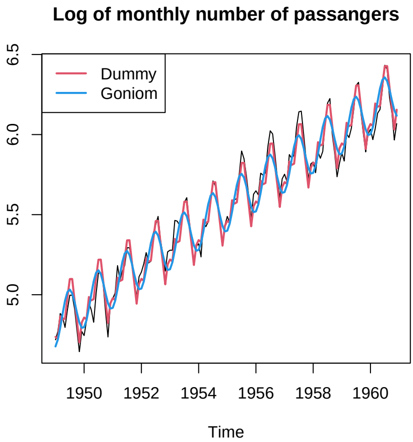
$$S_t = \beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right)$$

or

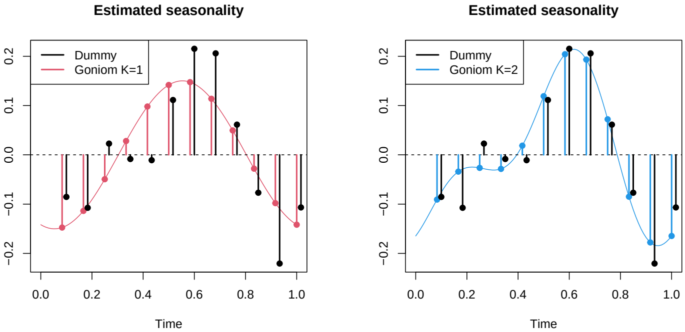
$$S_t = \beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right) + \beta_4 \sin\left(\frac{4\pi t}{12}\right) + \beta_5 \cos\left(\frac{4\pi t}{12}\right)$$

generally

$$S_t = \sum_{k=1}^{K} \left[\beta_{2k} \sin\left(\frac{2k\pi t}{12}\right) + \beta_{2k+1} \cos\left(\frac{2k\pi t}{12}\right)\right]$$



Log of monthly number of passangers
Dummy
Goniom
6.5
6.0
5.5
5.0
1950
1952
1954
1956
1958
1960
Time



Estimated seasonality
Dummy
Goniom K=1
0.2
0.1
0.0
−0.1
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
Time
Estimated seasonality
Dummy
Goniom K=2
0.2
0.1
0.0
−0.1
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
Time

# Holt-Winter's method

- extends Holt's method for seasonal data
- recursive formulas for level $L_t$, slope $T_t$ and seasonal component $I_t$
- parameters $\alpha, \gamma, \delta$

$\rightsquigarrow$ Read it in the book

# Cyclical component

Sometimes:

$$Y_t = Tr_t + S_t + C_t + E_t$$

$C_t$ cyclical component

- periodicities longer than one year

  e.g. several-year business cycle, cycles for hydrological series
- not clearly visible
- need to find number and length of corresponding periodicities

# Cyclical component

Sometimes:

$$Y_t = Tr_t + S_t + C_t + E_t$$

$C_t$ cyclical component

- periodicities longer than one year

  e.g. several-year business cycle, cycles for hydrological series
- not clearly visible
- need to find number and length of corresponding periodicities

$\rightsquigarrow$ tests of periodicity based on periodogram

# Cyclical component

Sometimes:

$$Y_t = Tr_t + S_t + C_t + E_t$$

$C_t$ cyclical component

- periodicities longer than one year

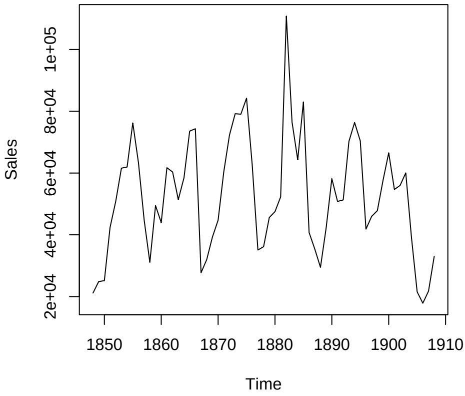
  e.g. several-year business cycle, cycles for hydrological series
- not clearly visible
- need to find number and length of corresponding periodicities

$\rightsquigarrow$ tests of periodicity based on periodogram

Assume that $\{Y_t\}$ already adjusted to trend and seasonality (if needed)

# Example

Yearly sales of mink fur coat (Hudson's Bay Company)



# Setting

Data $Y_1, \ldots, Y_n$

We want to decide if

$$Y_t = \mu + E_t$$

or

$$\begin{aligned} Y_t =& \mu + \sum_{i=1}^{k} [\alpha_i \cos(\omega_i t) + \beta_i \sin(\omega_i t)] + E_t \\ =& \mu + \sum_{i=1}^{k} \left[ \alpha_i \cos\left(\frac{2\pi t}{S_i} t\right) + \beta_i \sin\left(\frac{2\pi t}{S_i}\right)\right] + E_t \end{aligned}$$

# Setting

Data $Y_1, \ldots, Y_n$

We want to decide if

$$Y_t = \mu + E_t$$

or

$$\begin{aligned} Y_t =& \mu + \sum_{i=1}^{k} [\alpha_i \cos(\omega_i t) + \beta_i \sin(\omega_i t)] + E_t \\ =& \mu + \sum_{i=1}^{k} \left[ \alpha_i \cos\left( \frac{2\pi t}{S_i} t \right) + \beta_i \sin\left( \frac{2\pi t}{S_i} \right) \right] + E_t \end{aligned}$$

In that case we need to

- find $k$
- find $\omega_i$
- estimate $\alpha_i$, $\beta_i$

# Periodogram

$$I(\omega) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} Y_t e^{-it\omega}\right|^2, \quad \omega \in [0, \pi]$$

or equivalently

$$I(\omega) = \frac{1}{4\pi}\left(a^2(\omega) + b^2(\omega)\right),$$

where

$$a(\omega) = \sqrt{\frac{2}{n}}\sum_{t=1}^{n} Y_t \cos(\omega t), \quad b(\omega) = \sqrt{\frac{2}{n}}\sum_{t=1}^{n} Y_t \sin(\omega t).$$

See Stochastic processes II

# Periodogram

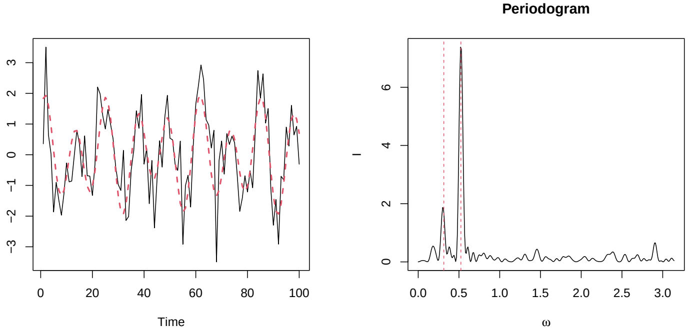
If

$$Y_t = \sum_{i=1}^{k} [\alpha_i \cos(\omega_i t) + \beta_i \sin(\omega_i t)] + E_t$$

then $I(\omega)$ has **local extremes** in $\omega_1, \ldots, \omega_k$

# Simulation

$$Y_t = \sin\left(\frac{2\pi}{12}t\right) + \cos\left(\frac{2\pi}{12}t\right) + \frac{1}{2}\sin\left(\frac{2\pi}{20}t\right) + \frac{1}{3}\cos\left(\frac{2\pi}{20}t\right) + E_t$$



# Detection of hidden periodicities

Data $Y_1, \ldots, Y_n$

- compute $I(\omega)$ for $Y_t - \overline{Y}_n$
- plot $I(\omega) \rightsquigarrow$ look for local extremes $\rightsquigarrow$ candidates for $\omega_j$

# Detection of hidden periodicities

Data $Y_1, \dots, Y_n$

- compute $I(\omega)$ for $Y_t - \overline{Y}_n$
- plot $I(\omega) \rightsquigarrow$ look for local extremes $\rightsquigarrow$ candidates for $\omega_j$
- **Fisher test**: For $j = 1, \dots, m = [(n-1)/2]$

$$W_j = \frac{I(\omega_j)}{\sum_{i=1}^m I(\omega_i)}, \quad \omega_j = \frac{2\pi j}{n}$$

and take

$$W = \max_{j=1,\dots,m} W_j$$

# Detection of hidden periodicities

Data $Y_1, \ldots, Y_n$

- compute $I(\omega)$ for $Y_t - \overline{Y}_n$
- plot $I(\omega) \rightsquigarrow$ look for local extremes $\rightsquigarrow$ candidates for $\omega_j$
- **Fisher test**: For $j = 1, \ldots, m = [(n-1)/2]$

$$W_j = \frac{I(\omega_j)}{\sum_{i=1}^m I(\omega_i)}, \quad \omega_j = \frac{2\pi j}{n}$$

  and take

$$W = \max_{j=1,\ldots,m} W_j$$

- Then under $H_0 : Y_t = E_t$ (no periodicity) for $E_t$ iid $\mathrm{N}(0, \sigma^2)$:

$$\mathrm{P}(W > w) = 1 - \sum_{i=0}^{\lfloor 1/w \rfloor} (-1)^i \binom{m}{i} (1 - iw)^{m-1}, \quad \text{pro } 0 < w < 1,$$

  $\rightsquigarrow$ critical value $g_{\alpha,m}$

# Detection of hidden periodicities

Data $Y_1, \ldots, Y_n$

- compute $I(\omega)$ for $Y_t - \overline{Y}_n$
- plot $I(\omega)$ $\rightsquigarrow$ look for local extremes $\rightsquigarrow$ candidates for $\omega_j$
- **Fisher test**: For $j = 1, \ldots, m = [(n-1)/2]$

$$W_j = \frac{I(\omega_j)}{\sum_{i=1}^{m} I(\omega_i)}, \quad \omega_j = \frac{2\pi j}{n}$$

  and take

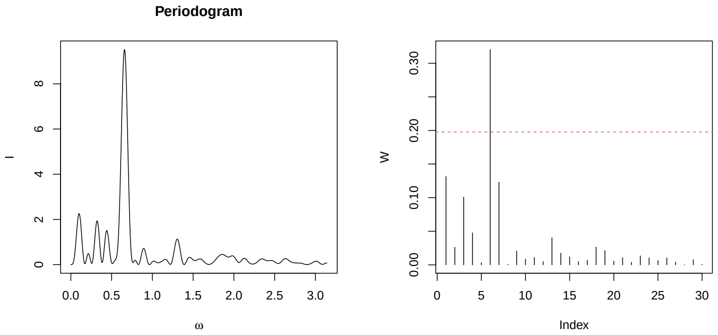
$$W = \max_{j=1,\ldots,m} W_j$$

- Then under $H_0 : Y_t = E_t$ (no periodicity) for $E_t$ iid $\mathrm{N}(0, \sigma^2)$:

$$\mathrm{P}(W > w) = 1 - \sum_{i=0}^{\lfloor 1/w \rfloor} (-1)^i \binom{m}{i} (1 - iw)^{m-1}, \quad \text{pro } 0 < w < 1,$$

  $\rightsquigarrow$ critical value $g_{\alpha,m}$
- if $W > g_{\alpha,m}$ $\rightsquigarrow$ reject $H_0$ and repeat the test without the detected frequency

# Mink coats




$$\rightsquigarrow \omega^* = \frac{2\pi 6}{n} = \frac{2\pi}{10.17} = 0.618$$

10 years periodicity

# Mink coats (cont.)

Estimate

$$Y_t = \mu + \alpha_1 \cos(2\pi t/10) + \beta_1 \sin(2\pi t/10) + E_t$$