Aalto University
School of Business

Time Series Analysis
Dummy Variables

Tomi Seppälä
Department of Information and Service Management

Introduction to Dummy Variables

Dummy variables (aka. Indicator variables or Binary variables)

Seasonality:

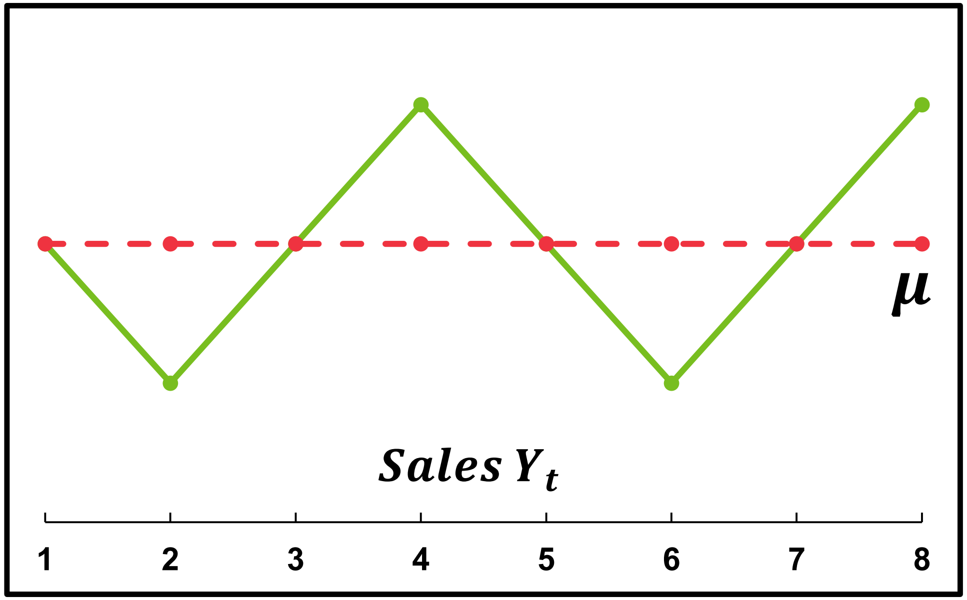
Example: Four (4) seasons 1, 2, 3, 4 (Quarters)

Time	Data quarter	D_1	D_2	D_3	D_4	Constant
1	1	1	0	0	0	1
2	2	0	1	0	0	1
3	3	0	0	1	0	1
4	4	0	0	0	1	1
5	1	1	0	0	0	1
6	2	0	1	0	0	1
7	3	0	0	1	0	1
8	4	0	0	0	1	1

↑

Categorical variable equivalent to 4 dummies (mutually exclusive)

Introduction to Dummy Variables



Model:

$$Y_t = \mu + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 D_3 + \gamma_4 D_4 + u_t$$

Complete linear dependence → We cannot estimate

$$\mu = D_1 + D_2 + D_3 + D_4$$

There are two alternatives

Introduction to Dummy Variables

Alternative 1:

$$Y = \mu + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 D_3 + u$$

- Drop one of the ***D***'s from the analysis
- Season 4 is the reference group (D_4) against which the other groups are compared
- Regression analysis table (ANOVA) gives t- and p-values compared to group 4:

$$H_0: \gamma_i = 0$$
$$H_1: \gamma_i \neq 0$$

separate individual tests for each $i = 1, 2, 3$

(Pairwise comparison between group ***i*** and 4)

Overall seasonality test: **is there seasonality in the data**:

Joint test $H_0: \gamma_1 = \gamma_2 = \gamma_3 = 0$ ← *No seasonality*

$$H_1: \gamma_1 \neq 0 \; or \; \gamma_2 \neq 0 \; or \; \gamma_3 \neq 0$$

ALWAYS DO OVERALL SEASONALITY TEST FIRST BEFORE ANY INDIVIDUAL TESTS!

Introduction to Dummy Variables

Alternative 2: Keep all the dummies, but drop μ

$$Y_t = \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 D_3 + \gamma_4 D_4 + u_t$$

No constant included

Joint test for overall seasonality (ANOVA to test means)

$$H_0: \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4$$
$$H_1: \gamma_i \neq \gamma_j \; for\ some\ i \neq j$$

Possible pairwise tests; for example:

$$H_0: \gamma_1 = \gamma_2$$
$$H_0: \gamma_1 \neq \gamma_2$$

Another way to deal with seasonality

$$Y_t = \mu + \alpha \cdot \sin(\mathrm{ct})$$

$$sin(0) = 0$$

$$sin(\frac{\pi}{2}) = 1$$

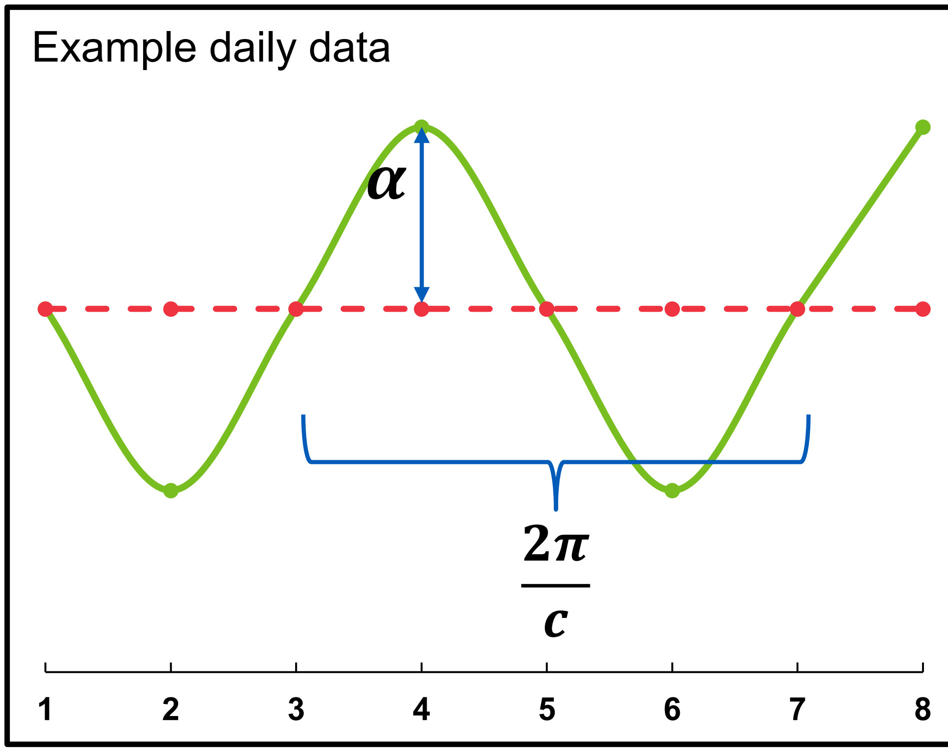
$$sin(\pi) = 0$$

$$sin\left(\frac{3\pi}{2}\right) = -1$$

$$sin(2\pi) = 0$$

Length of period:

$$\mathrm{ct} = 2\pi \rightarrow t = \frac{2\pi}{c}$$

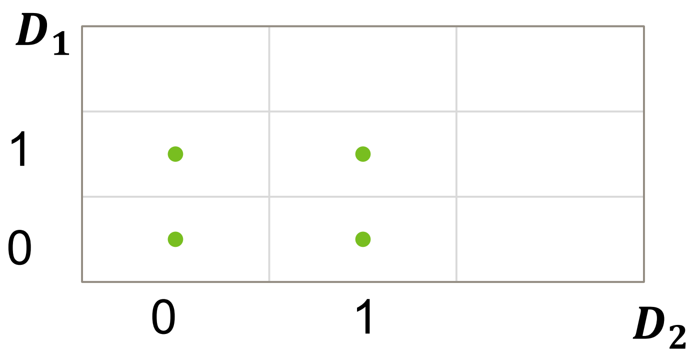


Interaction effects

Two separate dummy variables D_1, D_2, which could happen at the same time (or not). Note: these cannot be seasonal dummies, since they would be mutually exclusive

Interaction effect is $D_1 D_2$ (obtained with multiplying)

Four possibilities:



Model:

$$Y = \mu + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_{12} D_1 D_2 + u$$

Main effects ($\gamma_1 D_1 + \gamma_2 D_2$) *Interaction* ($\gamma_{12} D_1 D_2$)

Note: subscript t is left out for simplicity

Interaction effects

Model:

$$Y = \mu + \underbrace{\gamma_1 D_1 + \gamma_2 D_2}_{\textit{Main effects}} + \underbrace{\gamma_{12} D_1 D_2}_{\textit{Interaction}} + u$$

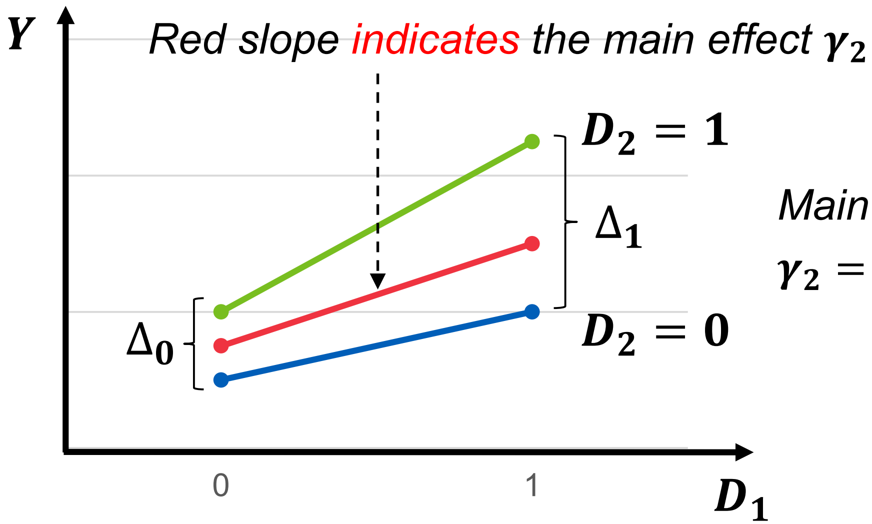
D_1 \ D_2	0	1
1	$\mu + \gamma_2 + u_t$	$\mu + \gamma_1 + \gamma_2 + \gamma_{12} + u_t$
0	$\mu + u_t$	$\mu + \gamma_1 + u_t$

Interaction effects

Test for interaction:

$$H_0: \gamma_{12} = 0$$
$$H_1: \gamma_{12} \neq 0$$

Test if the slopes are the same



Main effect
$$\gamma_2 = \frac{\Delta_1 + \Delta_0}{2}$$

Interpretation:

1. γ_{12} is the difference in slopes or alternatively
2. γ_{12} is the difference $\Delta_1 - \Delta_0$
3. If there is interaction, the effect on Y when changing D_1 from 0 to 1 is different for those with D_2=0 compared to those with D_2=1. i.e. green and blue slopes are the same