Modelling Time as a Circular Scale

HARVARD
School of Public Health

Department of Epidemiology
Miguel Angel Luque Fernandez,
Bizu Gelaye, Tyler Vander Weele, Hernandez-Diaz S, Michelle A. Williams,
Collaborators:
Ananth C.V, Qui C, Sanchez S.E, Cynthia Ferre, Anna Maria Siega-Riz,
Claudia Holzman, Daniel Enquobahrie, Nancy Dole

January 29, 2014

Table of contents

1 Chronobiology: Circular Time and Trigonometric Functions

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

24 hours (Circadian),

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

24 hours (Circadian),
30 days (Monthly),

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

  24 hours (Circadian),
  30 days (Monthly),
  Seasons (Seasonality),

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

  24 hours (Circadian),
  30 days (Monthly),
  Seasons (Seasonality),
  365 days (Annual), etc.

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

  24 hours (Circadian),
  30 days (Monthly),
  Seasons (Seasonality),
  365 days (Annual), etc.
- Cyclical events could be modeled as a time circular scale.

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

  24 hours (Circadian),
  30 days (Monthly),
  Seasons (Seasonality),
  365 days (Annual), etc.
- Cyclical events could be modeled as a **time circular scale**.

Chronobiology Definition and Time

Definition

Chronobiology is a discipline whose principles consider **time** as an essential dimension of biological phenomena.

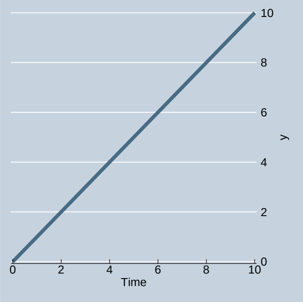
Time

- Biological time may be linear (chronological time) and cyclical (period time).
- **Cyclical time** could have several kinds of periodicities (biological rhythms).

  24 hours (Circadian),
  30 days (Monthly),
  Seasons (Seasonality),
  365 days (Annual), etc.
- Cyclical events could be modeled as a **time circular scale**.

Linear Time

Linear Time

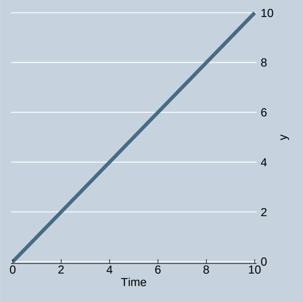


Linear Time

- Classical minimum squares regression modeling: Trend analysis.
- GLM: Rates, persons time at risk (Family Poisson, offset: time at risk and link log).

Linear Time

Linear Time

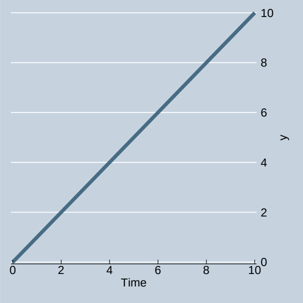


Linear Time

- Classical minimum squares regression modeling: Trend analysis.
- GLM: Rates, persons time at risk (Family Poisson, offset: time at risk and link log).
- Non functional shape assumption: Kaplan-Meier and Cox regression.

Linear Time

Linear Time

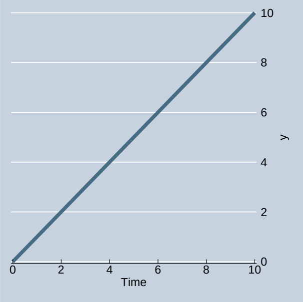


Linear Time

- Classical minimum squares regression modeling: Trend analysis.
- GLM: Rates, persons time at risk (Family Poisson, offset: time at risk and link log).
- Non functional shape assumption: Kaplan-Meier and Cox regression.

Linear Time

Linear Time

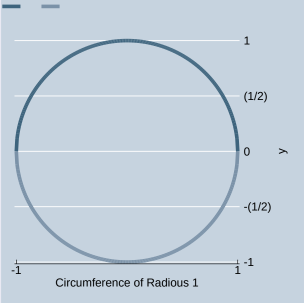


Linear Time

- Classical minimum squares regression modeling: Trend analysis.
- GLM: Rates, persons time at risk (Family Poisson, offset: time at risk and link log).
- Non functional shape assumption: Kaplan-Meier and Cox regression.

Circular Time

Circular Time

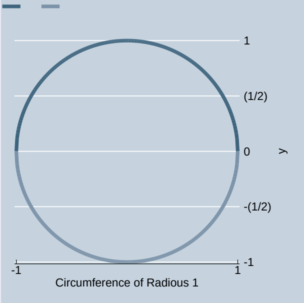


Circular time

- Assessing *periodicity*: Fourier Series (Periodogram).
- Describing *periodicity*: Data reduction.

Circular Time

Circular Time

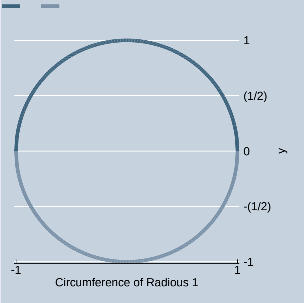


Circular time

- Assessing *periodicity*: Fourier Series (Periodogram).
- Describing *periodicity*: Data reduction.
- Modeling *periodicity*: Trigonometric predictors with sine and cosine terms (Trigonometric regression or cosinor model).

Circular Time

Circular Time



Circular time

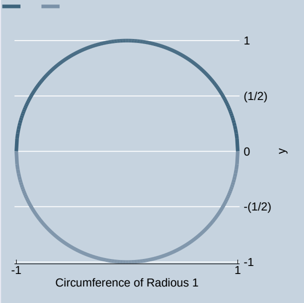
- Assessing *periodicity*: Fourier Series (Periodogram).
- Describing *periodicity*: Data reduction.
- Modeling *periodicity*: Trigonometric predictors with sine and cosine terms (Trigonometric regression or cosinor model).

Circular time modeling assumptions

Sinusoidal pattern

Circular Time

Circular Time



Circular time

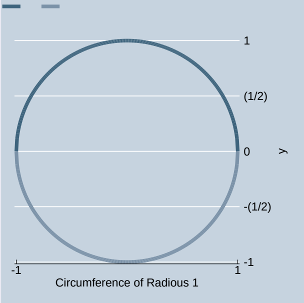
- Assessing *periodicity*: Fourier Series (Periodogram).
- Describing *periodicity*: Data reduction.
- Modeling *periodicity*: Trigonometric predictors with sine and cosine terms (Trigonometric regression or cosinor model).

Circular time modeling assumptions

Sinusoidal pattern — Stationary time series

Circular Time

Circular Time



Circular time

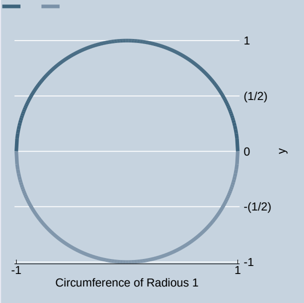
- Assessing *periodicity*: Fourier Series (Periodogram).
- Describing *periodicity*: Data reduction.
- Modeling *periodicity*: Trigonometric predictors with sine and cosine terms (Trigonometric regression or cosinor model).

Circular time modeling assumptions

Sinusoidal pattern — Stationary time series

Circular Time

Circular Time



Circular time

- Assessing *periodicity*: Fourier Series (Periodogram).
- Describing *periodicity*: Data reduction.
- Modeling *periodicity*: Trigonometric predictors with sine and cosine terms (Trigonometric regression or cosinor model).

Circular time modeling assumptions

Sinusoidal pattern — Stationary time series

Sine and Cosine

Sine and Cosine functions

- The steady rise-and-fall of the cosine and sine functions makes them ideal for modeling seasonality or circadian patterns.
- This rise-and-fall pattern is repeat.

Sine and Cosine

Sine and Cosine functions

- The steady rise-and-fall of the cosine and sine functions makes them ideal for modeling seasonality or circadian patterns.
- This rise-and-fall pattern is repeat.
- This repeating property of the sine and cosine functions means that we only need to consider times from 0 *to* $\leq 2\pi$.

Sine and Cosine

Sine and Cosine functions

- The steady rise-and-fall of the cosine and sine functions makes them ideal for modeling seasonality or circadian patterns.
- This rise-and-fall pattern is repeat.
- This repeating property of the sine and cosine functions means that we only need to consider times from 0 *to* $\leq 2\pi$.

Circular Time

The value of 2π is a key constant because it is the circumference of circle with radius 1.

Sine and Cosine

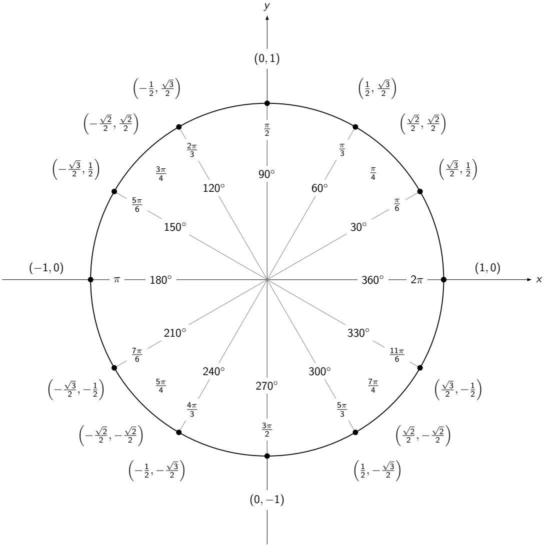
Sine and Cosine functions

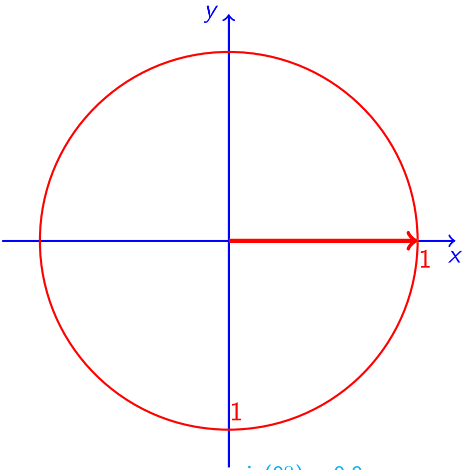
- The steady rise-and-fall of the cosine and sine functions makes them ideal for modeling seasonality or circadian patterns.
- This rise-and-fall pattern is repeat.
- This repeating property of the sine and cosine functions means that we only need to consider times from 0 *to* $\leq 2\pi$.

Circular Time

The value of 2π is a key constant because it is the circumference of circle with radius 1.

The trigonometric circle





$\sin(0^{o}) = 0.0$

$\cos(0^{o}) = 1.0$

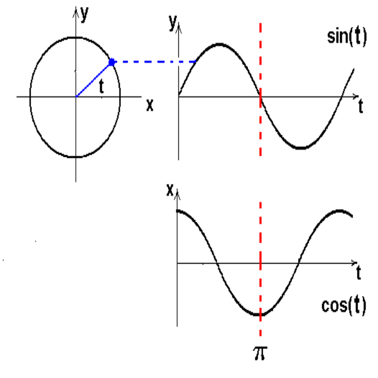
Circular Time

Trigonometric Functions

Link: Sum of Sine and Cosine

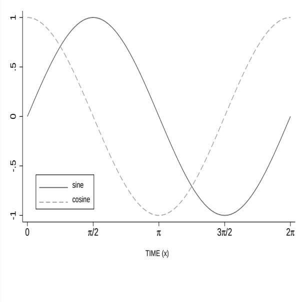
- Together the cosine and sine functions *can represent any point on the curve and the circle*.
- They are called *Trigonometric Functions*.
- The *rate of change* in cos(x) is given by sin(x) and vice versa.
- $\frac{d}{dx}cos(x) = -sin(x)$
- $\frac{d}{dx}sin(x) = cos(x)$

Sum of Sine and Cosine functions

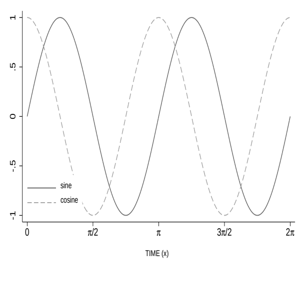


Sine and Cosine Functions

One cycle per 2π units of time

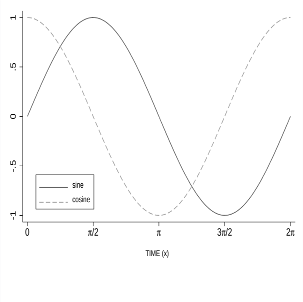


Two cycles per 2π units of time

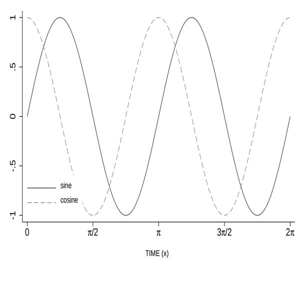


Sine and Cosine Functions

One cycle per 2π units of time



Two cycles per 2π units of time



2 Assessing a circular pattern

Table of contents

The examples used in this presentation: Work in progress

Modeling Vitamin D Serum Concentrations in a population of pregnant women.

- Data were drawn from an observational multicentric nested case-control study of 2,583 pregnant women using existing data and banked serum samples in the USA.
- Objective: To test the presence of a seasonal variation of 25OHD serum concentrations.
- We model maternal individual measurements of 25OHD serum concentrations (not repeat measurement within individuals).

Modeling the time of onset of Preterm Delivery

- Data were drawn from 476 women who delivered live births at three Hospitals in Lima, Peru, from January 2009 through July 2010.
- Objective: To model the time of onset of delivery in a sample of women who delivered a preterm infant.
- We model maternal self-reported time of onset of delivery.

The examples used in this presentation: Work in progress

Modeling Vitamin D Serum Concentrations in a population of pregnant women.

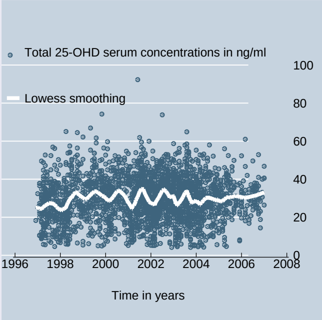
- Data were drawn from an observational multicentric nested case-control study of 2,583 pregnant women using existing data and banked serum samples in the USA.
- Objective: To test the presence of a seasonal variation of 25OHD serum concentrations.
- We model maternal individual measurements of 25OHD serum concentrations (not repeat measurement within individuals).

Modeling the time of onset of Preterm Delivery

- Data were drawn from 476 women who delivered live births at three Hospitals in Lima, Peru, from January 2009 through July 2010.
- Objective: To model the time of onset of delivery in a sample of women who delivered a preterm infant.
- We model maternal self-reported time of onset of delivery.

Assessing Seasonality

25OHD serum concentrations over 1996-2008: Lowess Smoothing, n= 2,583.



©MA Luque-Fernandez et al.Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

Assumptions

Assessing seasonality: First, Stationarity Time Series and Second a Sinusoidal or cyclic pattern (if modelled with a cosinor approach, it has to be symetric)

Fourier Time Series: Periodogram

Number of cycles in 2π time

- The periodogram $\mathrm{I}(w_j)$ is always positive, and it will be larger at frequencies that are strongly represented in the data.
- Therefore the number of time points needed to complete a cycle of 2π could be computed as the inverse of the Fourier frequency using:

$$1/f_j = \frac{2\pi}{w_j}$$

Formulae

$$I(w_j) = \frac{2}{n}(\hat{C}^2 + \hat{S}^2) \quad j = 1, \ldots n/2$$

$$\hat{C}^2 = 2\sum_{t=1}^{n} y_t \, cos(w_j t)/n,$$

$$\hat{S}^2 = 2\sum_{t=1}^{n} y_t \, sin(w_j t)/n,$$

Fourier Time Series: Periodogram

Number of cycles in 2π time

- The periodogram $I(w_j)$ is always positive, and it will be larger at frequencies that are strongly represented in the data.
- Therefore the number of time points needed to complete a cycle of 2π could be computed as the inverse of the Fourier frequency using:

$$1/f_j = \frac{2\pi}{w_j}$$

Formulae

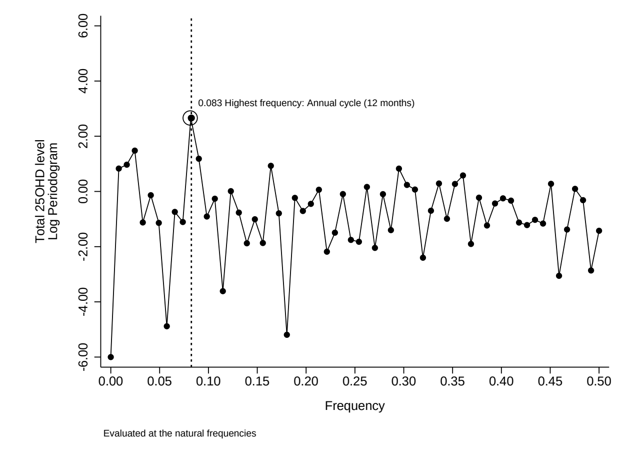
$$I(w_j) = \frac{2}{n}(\hat{C}^2 + \hat{S}^2) \quad j = 1, \ldots n/2$$

$$\hat{C}^2 = 2\sum_{t=1}^{n} y_t \, cos(w_j t)/n,$$

$$\hat{S}^2 = 2\sum_{t=1}^{n} y_t \, sin(w_j t)/n,$$

Example

Periodogram of 25OHD serum concentrations and highest frequency



©MA Luque-Fernandez et al.Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

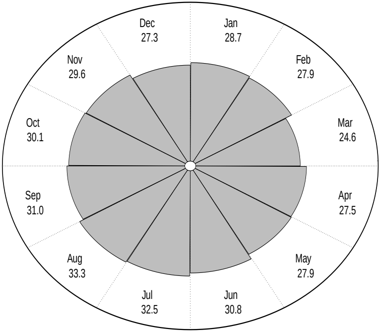
3 Describing Circadian and Seasonal Patterns

Table of contents

Data reduction

- Data reduction is one of the simplest methods for investigating a circadian, seasonal or annual pattern.
- A common method of data reduction is to group the data into 24 hours, 12 months, seasons, etc.
- Care needs to be taken when interpreting estimates, as they represent the average rates in each stratum.

Example:Circular Plot 25OHD serum concentrations

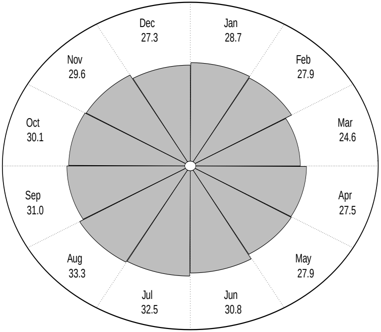


©MA Luque-Fernandez et al.Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic

Data reduction

- Data reduction is one of the simplest methods for investigating a circadian, seasonal or annual pattern.
- A common method of data reduction is to group the data into 24 hours, 12 months, seasons, etc.
- Care needs to be taken when interpreting estimates, as they represent the average rates in each stratum.

Example:Circular Plot 25OHD serum concentrations



©MA Luque-Fernandez et al.Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic

Grouping Data example

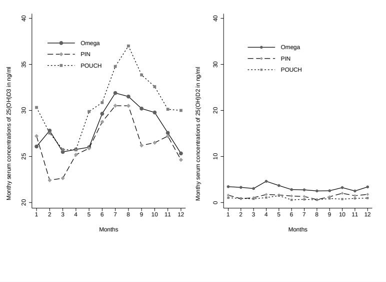
Grouping: tabular data

Mean and standard deviation of 25OHD serum concentrations by seasons, site and race, (n= 2,583).

	Black $\mu(\sigma^2)$,(n=649)			White $\mu(\sigma^2)$,(n=1934)		
	Omega(n=27)	Pin(n=350)	Pouch(n=272)	Omega(n=727)	Pin(n=642)	Pouch(n=565)
Winter	24.6(6.9)	17.5(8.6)	17.7(9.2)	29.7(8.4)	29.4(9.9)	34.6(10.9)
Spring	27.6(6.7)	18.0(8.8)	18.5(8.2)	29.4(8.9)	30.8(9.4)	33.5(10.3)
Summer	36.5(4.5)	21.6(8.5)	24.8(10.4)	33.4(8.6)	35.0(10.8)	39.3(9.5)
Fall	22.5(6.6)	19.4(9.8)	22.5(8.9)	31.9(7.7)	33.0(8.8)	36.7(10.6)
Annual	26.8(7.3)	19.0(9.0)	20.9(9.6)	31.2(8.6)	31.9(9.9)	36.1(10.5)

Grouping: Figure

Observed monthly means of 25OHD2 and D3 serum concentrations by site, (n= 2,583)



Grouping Data example

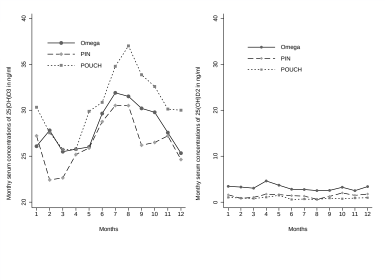
Grouping: tabular data

Mean and standard deviation of 25OHD serum concentrations by seasons, site and race, (n= 2,583).

	Black $\mu(\sigma^2)$,(n=649)			White $\mu(\sigma^2)$,(n=1934)		
	Omega(n=27)	Pin(n=350)	Pouch(n=272)	Omega(n=727)	Pin(n=642)	Pouch(n=565)
Winter	24.6(6.9)	17.5(8.6)	17.7(9.2)	29.7(8.4)	29.4(9.9)	34.6(10.9)
Spring	27.6(6.7)	18.0(8.8)	18.5(8.2)	29.4(8.9)	30.8(9.4)	33.5(10.3)
Summer	36.5(4.5)	21.6(8.5)	24.8(10.4)	33.4(8.6)	35.0(10.8)	39.3(9.5)
Fall	22.5(6.6)	19.4(9.8)	22.5(8.9)	31.9(7.7)	33.0(8.8)	36.7(10.6)
Annual	26.8(7.3)	19.0(9.0)	20.9(9.6)	31.2(8.6)	31.9(9.9)	36.1(10.5)

Grouping: Figure

Observed monthly means of 25OHD2 and D3 serum concentrations by site, (n= 2,583)



©MA Luque-Fernandez et al.Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

3 Modelling Stationary Circadian an Seasonal Patterns

Table of contents

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

- Time (months) is fitted as a categorical independent variable (x_i).

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

- Time (months) is fitted as a categorical independent variable (x_i).
- Usually we will use a Gaussian or a Poisson family with a link log.

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

- Time (months) is fitted as a categorical independent variable (x_i).
- Usually we will use a Gaussian or a Poisson family with a link log.
- A disadvantage is that it assumes complete independence between months.

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

- Time (months) is fitted as a categorical independent variable (x_i).
- Usually we will use a Gaussian or a Poisson family with a link log.
- A disadvantage is that it assumes complete independence between months.
- For many seasonal patterns this is unlikely to be true, as neighboring unit of time (months) are likely to be positively correlated.

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

- Time (months) is fitted as a categorical independent variable (x_i).
- Usually we will use a Gaussian or a Poisson family with a link log.
- A disadvantage is that it assumes complete independence between months.
- For many seasonal patterns this is unlikely to be true, as neighboring unit of time (months) are likely to be positively correlated.
- Although we can use Generalized Linear Mixed Models or a random intercept model.

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

- Time (months) is fitted as a categorical independent variable (x_i).
- Usually we will use a Gaussian or a Poisson family with a link log.
- A disadvantage is that it assumes complete independence between months.
- For many seasonal patterns this is unlikely to be true, as neighboring unit of time (months) are likely to be positively correlated.
- Although we can use Generalized Linear Mixed Models or a random intercept model.

Generalized Linear Models

GLM specification

$$y_i = \beta_0 + \beta_1 x_i \quad \text{where} \quad E(y) = \mu \quad \text{and} \quad \mu = X\beta$$

$$y_i \sim N(\mu_i, \sigma_i^2)$$

GLM

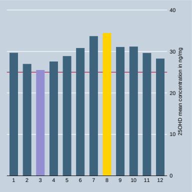
- Time (months) is fitted as a categorical independent variable (x_i).
- Usually we will use a Gaussian or a Poisson family with a link log.
- A disadvantage is that it assumes complete independence between months.
- For many seasonal patterns this is unlikely to be true, as neighboring unit of time (months) are likely to be positively correlated.
- Although we can use Generalized Linear Mixed Models or a random intercept model.

GLM Example

Observed monthly means, standard deviation and differences of 25OHD serum concentrations, (n= 2,583).

Month	Number of women tested	25(OH)D Mean	25(OH)D Std. Dev.	Absolute difference and 95%CI	Relative difference in percentage (%)
January	221	29.6	11.0	Ref.	Ref.
February	202	26.9	11.5	-2.68[(-4.83) to (-0.54)]	-9.1
March	233	25.5	10.8	-4.17[(-6.17) to (-2.16)]	-14.1
April	270	27.5	11.1	-2.11[(-4.07) to (-0.14)]	-7.1
May	241	28.8	10.7	-0.80[(-2.79) to 1.18]	-2.7
June	207	30.8	11.3	1.14 [(-0.97) to 3.25]	3.9
July	191	33.6	10.9	4.01 (1.89 to 6.13)	13.5
August	215	34.4	11.1	4.76 (2.68 to 6.84)	16.1
September	197	31.0	10.2	1.40 [(-0.63) to 3.44]	4.7
October	232	31.1	11.3	1.49 [(-0.57) to 3.54]	5.0
November	202	29.6	10.4	-0.04[(-2.08) to 1.99]	-0.1
December	172	28.2	11.0	-1.41[(-3.60) to 0.77]	-4.8

Figure. Observed monthly means of 25OHD serum concentrations, (n= 2,583)



©MA Luque-Fernandez et al. Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

Cosinor Model

Cosinor

The Cosinor model:

$$Y_t = c\,cos(w_t) + s\,sin(w_t)$$

t=1,...,n.

If we are interested in an annual seasonal cycle based on monthly data, then we would compute w_t as follow:

$$w_t = 2\pi\, f_t \quad \text{where}\,, \; f_t = \frac{month_t - 1}{12}$$

Amplitude and Phase

Where the Amplitude is:

$$A = \sqrt{c^2 + s^2}\,, \; (A \geq 0)$$

and the Phase [P(ϕ)]:

$$P = \begin{cases} arctan(s/c)\,, & c \geq 0, \\ arctan(s/c) + \pi\,, & c < 0\,, s \geq 0, \\ arctan(s/c) - \pi\,, & c < 0\,, s > 0. \end{cases}$$

To interpret the phase [P(ϕ)], it is preferable to transform this to a time scale using $P' = 12(P/2\pi) + 1$ for monthly data.

Cosinor Model

Cosinor

The Cosinor model:

$$Y_t = c\,cos(w_t) + s\,sin(w_t)$$

t=1,...,n.

If we are interested in an annual seasonal cycle based on monthly data, then we would compute w_t as follow:

$$w_t = 2\pi\, f_t \quad \text{where}\ ,\ f_t = \frac{month_t - 1}{12}$$

Amplitude and Phase

Where the Amplitude is:

$$A = \sqrt{c^2 + s^2}\ ,\ (A \geq 0)$$

and the Phase [P(ϕ)]:

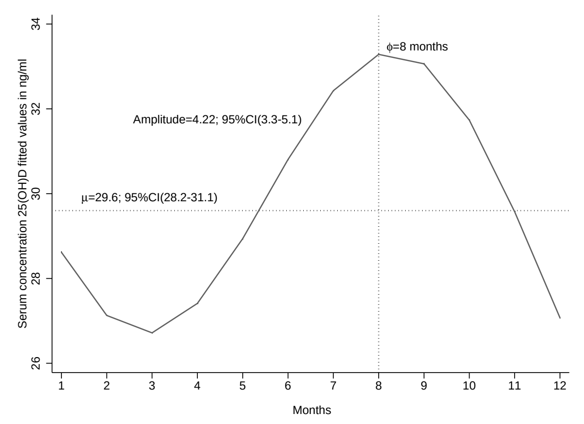
$$P = \begin{cases} arctan(s/c)\ ,\ c \geq 0, \\ arctan(s/c) + \pi\ ,\ c < 0\ , s \geq 0, \\ arctan(s/c) - \pi\ ,\ c < 0\ , s > 0. \end{cases}$$

To interpret the phase [P(ϕ)], it is preferable to transform this to a time scale using $P' = 12(P/2\pi) + 1$ for monthly data.

Cosinor Modelling Example

Fitted Means: Univariate example

Modelled distribution of 25OHD serum concentrations, n=2,583



©MA Luque-Fernandez et al. Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

Cosinor Inference Example

Crude and Adjusted Annual Means of 25OHD and Mean Peak-Trough Difference in 25OHD (n= 2,583)

Variables		Crude 25(OH)D Annual Mean, ng/mL	95%CI	Adjusted 25(OH)D Annual Mean, ng/mL[b]	95%CI	25(OH)D Mean Peak-Trough difference, ng/mL	95%CI
Maternal Age							
	15-24	27.7	27.1, 28.4	28.2	27.4, 28.9	5.9	4.9, 7.0
	25-34	29.9	29.4, 30.3	30.4	29.9, 30.9	8.2	6.8, 9.6
	$\geq$35	31.9	31.2, 32.8	29.7	28.9, 30.5	7.5	5.9, 9.0
	P for difference	<0.001		0.003		0.005	
Race							
	Black	20.2	19.5, 21.0	19.6	18.9, 20.4	5.9	4.9, 7.0
	White	32.8	32.4, 33,2	33.0	32.6, 33.5	7.1	5.6, 8.6
	P for difference	<0.001		<0.001		<0.001	
Site							
	Omega (Seattle)	30.8	30.0, 31.6	30.9	30.3, 31.5	5.7	4.7, 6.7
	Pin (North Carolina)	27.5	26.8, 28.2	27.5	26.8, 28.1	2.3	0.9, 3.7
	Pouch (Michigan)	31.2	30.4, 31.9	31.2	30.5, 31.8	6.0	4.6, 7.4
	P for difference	0.372		0.236		0.001	
Gestational week							
	I Trimester	27.3	25.7, 28.8	26.8	28.8, 27.9	5.9	4.9, 7.0
	II Trimester	29.8	29.4, 30.3	29.8	29.5, 30.3	8.9	7.4, 10.5
	P for difference	0.002		<0.001		0.001	
Maternal Education							
	Highschool or less	26.7	26.0, 27.5	28.0	27.2, 28.8	5.9	4.9, 7.0
	Post Highschool	30.9	30.4, 31.5	30.4	29.9, 30.9	8.3	6.9, 9.7
	P for difference	<0.001		<0.001		0.001	
Pre-pregnancy BMI in kg/m^2							
	<25	31.4	30.8, 31.9	30.9	30.4, 31.4	5.9	4.9, 7.0
	25-30	29.4	28.5, 30.3	29.4	28.6, 30.3	4.5	3.1, 6.0
	>30	25.1	24.2, 26.0	26.5	25.6, 27.3	1.6	0.1, 3.0
	P for difference	<0.001		<0.001		<0.001	

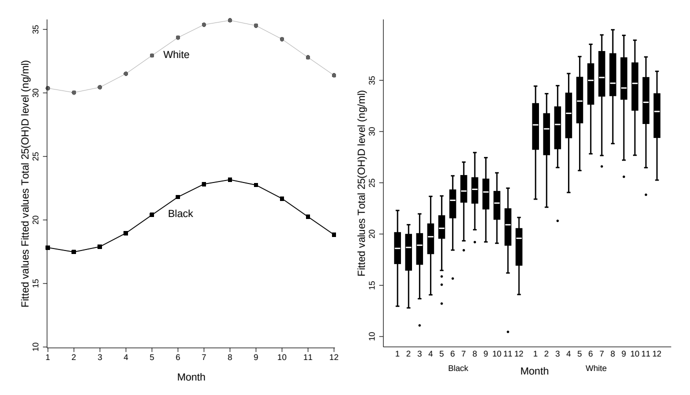
[a] Models were adjusted for the main effect of maternal age, gestational weeks, race and study site.
[b] Annual means were centered to reflect study population values for maternal age, gestational weeks, race and study site.

©MA Luque-Fernandez et al.Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

Cosinor Modelling Example

Fitted Means: Bivariate example

Distribution of 25OHD serum concentrations modelled with a bivariate Stationary Cosinor Model by race (n= 2,583)



©MA Luque-Fernandez et al. Seasonal Variation of 25-Hydroxyvitamin D among non-Hispanic Black and White Pregnant Women from Three US Pregnancy Cohorts. Pediatrics and Perinatal Epidemiology 2013

Cubic Splines

A cubic spline function with K knots is given by:

$$f(x) = \sum_{j=0}^{3} \beta_{0j} x^j + \sum_{l=1}^{k} \beta_i (x - t_l)^3 +,$$

where $t_l, l = 1, ..., k$ are the k knots. And x is related with the outcome as:

$$y_i = f(x_i) + \epsilon_i$$

Number of knots

Choosing the knots

- Knots are usually placed at quantiles of the data or at regularly spaced intervals.
- Choosing the number, rather than the placement, seems to be more crucial to the fit.

Number of knots

Choosing the knots

- Knots are usually placed at quantiles of the data or at regularly spaced intervals.
- Choosing the number, rather than the placement, seems to be more crucial to the fit.
- It is better to choose a number of knots that represents the curvature you believe to be present in the data.

Number of knots

Choosing the knots

- Knots are usually placed at quantiles of the data or at regularly spaced intervals.
- Choosing the number, rather than the placement, seems to be more crucial to the fit.
- It is better to choose a number of knots that represents the curvature you believe to be present in the data.
- Also the knots could be placed at points in the data where you expect significant changes in the relationship between the predictor and the outcome to occur.

Number of knots

Choosing the knots

- Knots are usually placed at quantiles of the data or at regularly spaced intervals.
- Choosing the number, rather than the placement, seems to be more crucial to the fit.
- It is better to choose a number of knots that represents the curvature you believe to be present in the data.
- Also the knots could be placed at points in the data where you expect significant changes in the relationship between the predictor and the outcome to occur.

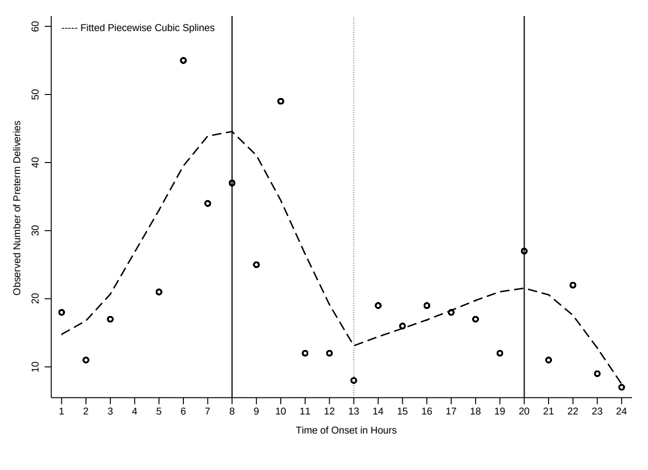
Number of knots

Choosing the knots

- Knots are usually placed at quantiles of the data or at regularly spaced intervals.
- Choosing the number, rather than the placement, seems to be more crucial to the fit.
- It is better to choose a number of knots that represents the curvature you believe to be present in the data.
- Also the knots could be placed at points in the data where you expect significant changes in the relationship between the predictor and the outcome to occur.

Cubic Spline Example

Modelling the Onset of labor in a sample of Preterm Delivery infants, n= 476



References

Some important references

1. A. G. Barnett, A. J. Dobson, E. Library., Analysing seasonal health data, Springer, 2010.
2. N. I. Fisher, Statistical analysis of circular data, Cambridge University Press, Cambridge, 1993.
3. P. Bloomfield, Fourier analysis of time series: an introduction, 2nd Edition, Wiley series in probability and statistics Applied probability and statistics, Wiley, New York ; Chichester, 2000.
4. T. W. Korner, Fourier analysis, Cambridge University Press, Cambridge, 1988
5. J. W. Hardin, J. Hilbe, Generalized linear models and extensions, 3rd Edition, Stata Press, College Station, Tex., 2012.