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**
   - Definition
   - Time
   - Circular Time: Sine and Cosine Functions
   - Circular Time: Sine and Cosine Functions

2. **Assessing a circular pattern**
   - The examples used in this presentation: Work in progress
   - Time plot
   - Periodogram

3. **Describing Circadian and Seasonal Patterns**
   - Grouping Data

4. **Modelling Stationary Circadian an Seasonal Patterns**
   - GLMs
   - Cosionor Model
   - Cubic Splines

5. **References**
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 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),
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**.

---

**Circular Time: Sine and Cosine Functions**

**Modelling Stationary Circadian and Seasonal Patterns**

**References**
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

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

- 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

**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

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

- **Definition**: Time as a circular scale.
- **Assessing a circular pattern**
- **Describing Circadian and Seasonal Patterns**
- **Modelling Stationary Circadian and Seasonal Patterns**
- **References**
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

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**

- **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

- **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
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 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 $2\pi$ is a key constant because it is the circumference of a circle with radius 1.
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 $2\pi$ is a key constant because it is the circumference of a circle with radius 1.
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 $0 \leq 2\pi$.

Circular Time

The value of $2\pi$ is a key constant because it is the circumference of circle with radius 1.
The trigonometric circle

The trigonometric circle is a representation of circular time, where angles are measured in radians from the positive x-axis. The circle is divided into sections corresponding to the values of sine and cosine functions at various angles. Points on the circle correspond to specific angles and their associated sine and cosine values.

- The x-axis represents the cosine function, and the y-axis represents the sine function.
- Angles are measured in radians from the positive x-axis, with key points marked at 0, π/2, π, 3π/2, and 2π radians.
- The circle is divided into 360°, with each 15° marking a significant angle.
- Key points on the circle include:
  - (0, 1): (0, 0)
  - (1, 0)
  - (0, -1)
  - (-1, 0)
  - Coordinates corresponding to the angles are shown, such as (1/2, √3/2) for π/6 radians.
  - Other significant points include (0, -√3), (0, 1), (1, 0), and (-1, 0).

This circle is a fundamental tool in understanding circadian and seasonal patterns in chronobiology.
Chronobiology
Assessing a circular pattern
Describing Circadian and Seasonal Patterns
Modelling Stationary Circadian and Seasonal Patterns
References

Definition
Time
Circular Time: Sine and Cosine Functions
Circular Time: Sine and Cosine Functions

Modelling Time as a Circular Scale
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.
  \[
  \frac{d}{dx} \cos(x) = -\sin(x)
  \]
  \[
  \frac{d}{dx} \sin(x) = \cos(x)
  \]
Sine and Cosine Functions

One cycle per $2\pi$ units of time

Two cycles per $2\pi$ units of time

©MA Luque-Fernandez
Sine and Cosine Functions

One cycle per $2\pi$ units of time

Two cycles per $2\pi$ units of time

©MA Luque-Fernandez
2 Assessing a circular pattern
Table of contents

1 Chronobiology
   • Definition
   • Time
   • Circular Time: Sine and Cosine Functions
   • Circular Time: Sine and Cosine Functions

2 Assessing a circular pattern
   • The examples used in this presentation: Work in progress
   • Time plot
   • Periodogram

3 Describing Circadian and Seasonal Patterns
   • Grouping Data

4 Modelling Stationary Circadian and Seasonal Patterns
   • GLMs
   • Cosinor Model
   • Cubic Splines

5 References
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


Assumptions

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

Number of cycles in $2\pi$ 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 $2\pi$ 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 $2\pi$ 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 $2\pi$ 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

3 Describing Circadian and Seasonal Patterns
Table of contents

1 Chronobiology
   • Definition
   • Time
   • Circular Time: Sine and Cosine Functions

2 Assessing a circular pattern
   • The examples used in this presentation: Work in progress
   • Time plot
   • Periodogram

3 Describing Circadian and Seasonal Patterns
   • Grouping Data

4 Modelling Stationary Circadian an Seasonal Patterns
   • GLMs
   • Cosionor Model
   • Cubic Splines

5 References
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
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
Grouping Data example

**Grouping: tabular data**

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

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

**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, \( n = 2,583 \).

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

Grouping: Figure

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

3 Modelling Stationary Circadian and Seasonal Patterns
# Table of contents

1. **Chronobiology**
   - Definition
   - Time
   - Circular Time: Sine and Cosine Functions
   - Circular Time: Sine and Cosine Functions

2. **Assessing a circular pattern**
   - The examples used in this presentation: Work in progress
   - Time plot
   - Periodogram

3. **Describing Circadian and Seasonal Patterns**
   - Grouping Data

4. **Modelling Stationary Circadian an Seasonal Patterns**
   - GLMs
   - Cosionor Model
   - Cubic Splines

5. **References**
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

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

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

Cosinor

The Cosinor model:

\[ Y_t = c \cos(w_t) + s \sin(w_t) \]

where \( t = 1, \ldots, n \).

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

\[ w_t = 2\pi f_t \]

where \( f_t = \frac{\text{month}_t - 1}{12} \)

Amplitude and Phase

Where the Amplitude is:

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

and the Phase \( [P(\phi)] \):

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

To interpret the phase \( [P(\phi)] \), 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:

\[ w_t = 2\pi f_t \quad \text{where} \quad f_t = \frac{\text{month}_t - 1}{12} \]

**Amplitude and Phase**

Where the Amplitude is:

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

and the Phase \([P(\phi)]\):

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

To interpret the phase \([P(\phi)]\), 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

Serum concentration 25(OH)D fitted values in ng/ml

## Cosinor Inference Example

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

<table>
<thead>
<tr>
<th>Variables</th>
<th>Crude 25(OH)D</th>
<th>Adjusted 25(OH)D</th>
<th>25(OH)D Mean Peak-Trough Difference, ng/mL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual Mean, ng/mL</td>
<td>95%CI</td>
<td>Annual Mean, ng/mL</td>
</tr>
<tr>
<td>Maternal Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-24</td>
<td>27.7</td>
<td>27.1, 28.4</td>
<td>28.2</td>
</tr>
<tr>
<td>25-34</td>
<td>29.9</td>
<td>29.4, 30.3</td>
<td>30.4</td>
</tr>
<tr>
<td>≥35</td>
<td>31.9</td>
<td>31.2, 32.8</td>
<td>29.7</td>
</tr>
<tr>
<td>P for difference</td>
<td>&lt;0.001</td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>20.2</td>
<td>19.5, 21.0</td>
<td>19.6</td>
</tr>
<tr>
<td>White</td>
<td>32.8</td>
<td>32.4, 33.2</td>
<td>33.0</td>
</tr>
<tr>
<td>P for difference</td>
<td>&lt;0.001</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omega (Seattle)</td>
<td>30.8</td>
<td>30.0, 31.6</td>
<td>30.9</td>
</tr>
<tr>
<td>Pin (North Carolina)</td>
<td>27.5</td>
<td>26.8, 28.2</td>
<td>27.5</td>
</tr>
<tr>
<td>Pouch (Michigan)</td>
<td>31.2</td>
<td>30.4, 31.9</td>
<td>31.2</td>
</tr>
<tr>
<td>P for difference</td>
<td>0.372</td>
<td>0.236</td>
<td>0.003</td>
</tr>
<tr>
<td>Gestational week</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I Trimester</td>
<td>27.3</td>
<td>25.7, 28.8</td>
<td>26.8</td>
</tr>
<tr>
<td>II Trimester</td>
<td>29.8</td>
<td>29.4, 30.3</td>
<td>29.8</td>
</tr>
<tr>
<td>P for difference</td>
<td>0.002</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Maternal Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highschool or less</td>
<td>26.7</td>
<td>26.0, 27.5</td>
<td>28.0</td>
</tr>
<tr>
<td>Post Highschool</td>
<td>30.9</td>
<td>30.4, 31.5</td>
<td>30.4</td>
</tr>
<tr>
<td>P for difference</td>
<td>&lt;0.001</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pre-pregnancy BMI in kg/m²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>31.4</td>
<td>30.8, 31.9</td>
<td>30.9</td>
</tr>
<tr>
<td>25-30</td>
<td>29.4</td>
<td>28.5, 30.3</td>
<td>29.4</td>
</tr>
<tr>
<td>&gt;30</td>
<td>25.1</td>
<td>24.2, 26.0</td>
<td>26.5</td>
</tr>
<tr>
<td>P for difference</td>
<td>&lt;0.001</td>
<td></td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

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.

Cosinor Modelling Example

**Fitted Means: Bivariate example**

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

A cubic spline function with \( K \) knots is given by:

\[
f(x) = \sum_{j=0}^{3} \beta_0 j^j + \sum_{l=1}^{k} \beta_l (x - t_l)^3 + ,
\]

where \( t_l, l = 1, \ldots, k \) 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.
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.
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.
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

©MA Luque-Fernandez et al., non published data

HSPH-Department of Epidemiology | Modelling Time as a Circular Scale
### References

<table>
<thead>
<tr>
<th>Reference</th>
<th>Details</th>
</tr>
</thead>
</table>