An Introduction to Circular Regression

Based on "Circular Regression" by T.D.Downs and K.V.Mardia

Hengrui Luo and Kori Khan

Department of Statistics
The Ohio State University

November 25, 2018

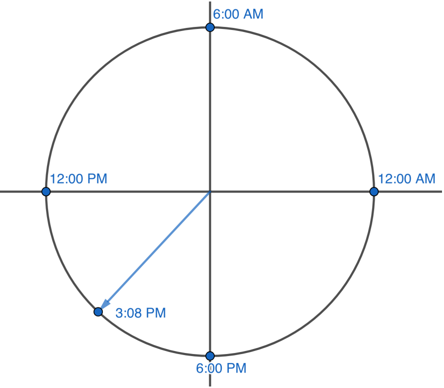
A Motivating Example

- Consider the case in which you have observed the peak systolic blood pressure for two different time periods: $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$

	Person 1	Person 2	...	Person 10
$\boldsymbol{S}_1$	5:12 p.m.	10:52 p.m.	...	9:04 p.m.
$\boldsymbol{S}_2$	5:28 p.m.	11:04 p.m.	...	9:52 p.m.

- You are interested in cyclic variations in this circadian process (i.e., are the peak times the same for the two time periods)
- Approaches in time series may make assumptions that are not actually met

A Motivating Example



- An alternative approach might be to represent $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ as circular variables or random angles on $\mathbb{S}^1$
- We can treat each hour as 15° on $\mathbb{S}^1$
- In the GLM setting, we might then answer the question with $g(E(\boldsymbol{S}_1)) = \beta \boldsymbol{S}_2$

The von Mises Distribution

- We say that a random angle v has the von Mises distribution with mean μ and nonnegative concentration parameter κ when the density for u is:

$$f_{\mu,\kappa}(v) = \frac{1}{2\pi I_0(\kappa)} \exp\left(\kappa \cos(v-\mu)\right)$$

  - μ is the direction from the origin to the center of gravity.
  - $\rho = \frac{I_1(\kappa)}{I_0(\kappa)}$ is the distance from the origin to the center of gravity.
  - Expectation $\mathrm{E}(\cos t, \sin t) = \rho(\cos\mu, \sin\mu)$.
  - Rotational invariance $t \sim \mathrm{M}(\mu,\kappa) \Rightarrow t+\theta \sim \mathrm{M}(\mu+\theta,\kappa)$, for $\theta \in (-\pi,\pi]$

Circular Regression

- We assume that u is an angular variable

$$\textbf{Distribution } v \mid u \sim \mathrm{M}(\mu, \kappa)$$

$$\textbf{Link Function } \tan \frac{1}{2}(\mu - \beta) = \omega \ \tan \frac{1}{2}(u - \alpha)$$

$$\implies \mu = \beta + 2\arctan\{\omega \ \tan \frac{1}{2}(u - \alpha)\}$$

  - α, β are angular location parameters
  - $\omega \in [-1, 1]$ is a slope parameter
  - **Angular Error:** $v - \mu \sim M(0, \kappa)$

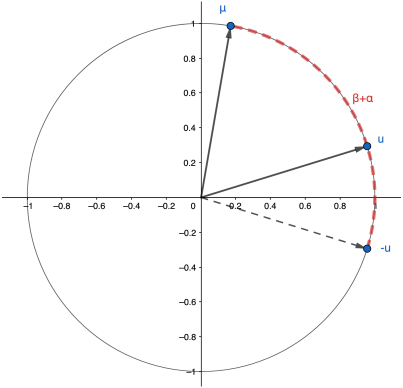
Circular Regression

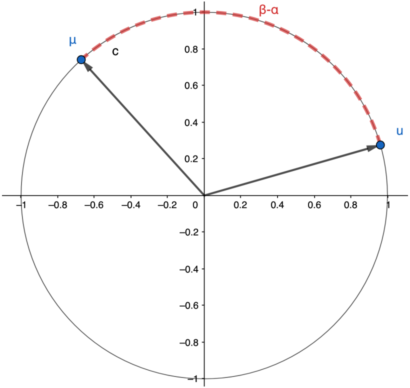
- To develop intuition, assume that $\omega \in \{-1, 0, 1\}$

$$\textbf{Distribution } v \mid u \sim \mathrm{M}(\mu, \kappa)$$
$$\textbf{Link Function } \tan\frac{1}{2}(\mu - \beta) = \omega \ \tan\frac{1}{2}(u - \alpha)$$
$$\implies \mu = \omega u + (\beta - \omega\alpha)$$

  - $\omega = 0 \implies \mu = \beta$, (v and u are independent)
  - $\omega = 1 \implies \mu$ is a rotation of u through $\beta - \alpha$
  - $\omega = -1 \implies \mu$ is a rotation of $-u$ through $\beta + \alpha$
- Rotational Models are a special case of more general framework

Circular Regression





Classification of Models.

- Let $\Theta = (\alpha, \beta, \omega)$ be the parameter space, we can classify models based on the functional relationships between the parameters

  - **Class A**: Independent modeling

$$\Longrightarrow \alpha, \beta, \omega \text{ are functionally independent}$$

  - **Class B**: Symmetry modeling

$$\Longrightarrow \alpha \pm \beta = 0 \Leftrightarrow (\alpha, \beta, \omega) = (\alpha, \pm\alpha, \omega)$$

  - **Class C**: Special orientation $\omega = 0, \pm 1$ and presumed known

$\Longrightarrow \beta - \omega\alpha$ is parameter of interest

- Why this classification?
  - We need to perform testing, $\Theta_C \subset \Theta_B \subset \Theta_A$
  - Helps in development of methods to obtain estimates

Statistical Inference

- Assume that we want to test whether $\boldsymbol{v} = (v_1, \ldots, v_n)$ and $\boldsymbol{u} = (u_1, \ldots, u_n)$ are independent:

$$H_0 : \omega = 0$$

$$H_1 : \omega \neq 0$$

- Note the alternative hypothesis must contain the parameter space under the null

- When ρ and n are sufficiently large, then under the null hypothesis

$$\frac{(n-3)(\hat{\rho}_A - \hat{\rho}_{C,0})}{2(1-\hat{\rho}_A)} \sim F_{2,n-3}$$

  - Where $\hat{\rho}_A$ is the MLE estimate for ρ under Class A models and $\hat{\rho}_{C,0}$ is the MLE estimate for ρ under Class C models with $\omega = 0$

Statistical Inference

- Returning to blood pressure example, with $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ peak systolic blood pressure from two different time periods.

	Person 1	Person 2	...	Person 10
$\boldsymbol{S}_1$	-102	-17	...	-44
$\boldsymbol{S}_2$	-98	-14	...	-32

- Consider the set-up:

$$\textbf{Distribution } \boldsymbol{S}_1 \mid \boldsymbol{S}_2 \sim \mathrm{M}(\mu, \kappa)$$

$$\textbf{Link Function } \tan\frac{1}{2}(\mu - \beta) = \omega \ \tan\frac{1}{2}(u - \alpha)$$

$$\implies \mu = \beta + 2\arctan\{\omega \ \tan\frac{1}{2}(\boldsymbol{S}_2 - \alpha)\}$$

$$\mu = \beta + 2\arctan\{\omega \ \tan\frac{1}{2}(\boldsymbol{S}_2 - \alpha)\}$$

- Question 1: Is there any association between $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$?
  - $H_0 : \omega = 0$ vs $H_0 : \omega \neq 0$
  - $\frac{(10-3)(\hat{\rho}_A - \hat{\rho}_{C,0})}{(3-1)(1-\hat{\rho}_A)} = 30.5 \sim F_{3-1,10-3}$
- Question 2: Are the blood pressure peak times identical for $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$?

$$H_0 : \omega = 1 \ \& \ \beta - \alpha = 0 \text{ vs } H_1 : \omega \neq 1 \text{ or } \beta - \alpha \neq 0$$

$$\frac{(10-3)(\hat{\rho}_A - \hat{\rho}_{C,1,0})}{3(1-\hat{\rho}_A)} = .724 \sim F_{3-0,10-3}$$

For Further Reading I

- Downs, Thomas D., and K. V. Mardia. "Circular regression." Biometrika 89.3 (2002): 683-698.
- Kato, Shogo, Kunio Shimizu, and Grace S. Shieh. "A circular–circular regression model." Statistica Sinica (2008): 633-645.
- Chernov, Nikolai. Circular and linear regression: Fitting circles and lines by least squares. CRC Press, 2010.