Savitribai Phule Pune University

सावित्रीबाई फुले पुणे विद्यापीठ

Department of Statistics

Savitribai Phule Pune University

MSC-II Project

Detection of Circadian rhythmicity using Statistical techniques

-Ayyaj Shaikh(1646)
-Amol Markad(1630)
-Priyanka Dandale(1614)

Project Guide: Dr. Akanksha S. Kashikar

Savitribai Phule Pune University

सावित्रीबाई फुले पुणे विद्यापीठ

Certificate

This is to certify that,

Mr. Ayyaj Shaikh (PRN 1646)
Mr. Amol Markad (PRN 1630)
Ms. Priyanka Dandale (PRN 1614)

of M.Sc.-II (Statistics) have successfully completed their project on "Detection of Circadian rhythmicity using Statistical techniques " under the guidance of Dr. Akanksha Kashikar and have submitted this project report as a part of the course ST-403, for the degree of M.Sc. (Statistics) from Savitribai Phule Pune University, in the academic year 2017-2018.

DATE:

PLACE: PUNE

Dr. Akanksha Kashikar
Project Guide
Department of Statistics,
Savitribai Phule Pune University

Prof. T. V. Ramanathan
Head of Department
Department of Statistics,
Savitribai Phule Pune University

Acknowledgement

We would like to extend our sincere thanks to all who supported us in this project. We are highly indebted to Dr. Akanksha Kashikar, Department of Statistics, Savitribai Phule Pune University, for her guidance and constant supervision as well as for providing necessary information regarding the project and also for her support in completing the project.

A special gratitude and thanks to Prof. T. V. Ramanathan, Head of the department, Department of Statistics, Savitribai Phule Pune University, for providing us various facilities at the department to complete the project.

Our thanks to the entire faculty and research students of Department of Statistics, Savitribai Phule Pune University for willingly helping us during this work.

Finally we would like to express our gratitude towards our family members and friends for encouragement and immense support which helped us in completion of this project.

Contents

1. Introduction

Certain prerequisites are required in order to understand the purpose of our study.

1.1 What is Circadian rhythm?

A cycle or rhythm is present when a phenomenon repeatedly goes through a peak and a trough over a specific period of time.

A circadian rhythm is a roughly 24 hour cycle in the physiological processes of living beings, including plants, animals and fungi.

There are many examples of circadian rhythms, such as the sleep-wake cycle, the body-temperature cycle, and the cycles in which a number of hormones are secreted.

1.2 Terminology

1.2.1 Period:

It is the duration of a full cycle (that is, the time distance between one peak and the next).

1.2.2 MESOR:

It stands for Midline Estimating Statistic Of Rhythm .

- The mean level is the central value around which the oscillation occurs.

- It is a rhythm-adjusted mean that differs from the arithmetic mean when the data are not equidistant and/or do not covers an integer number of cycles.

1.2.3 Amplitude:

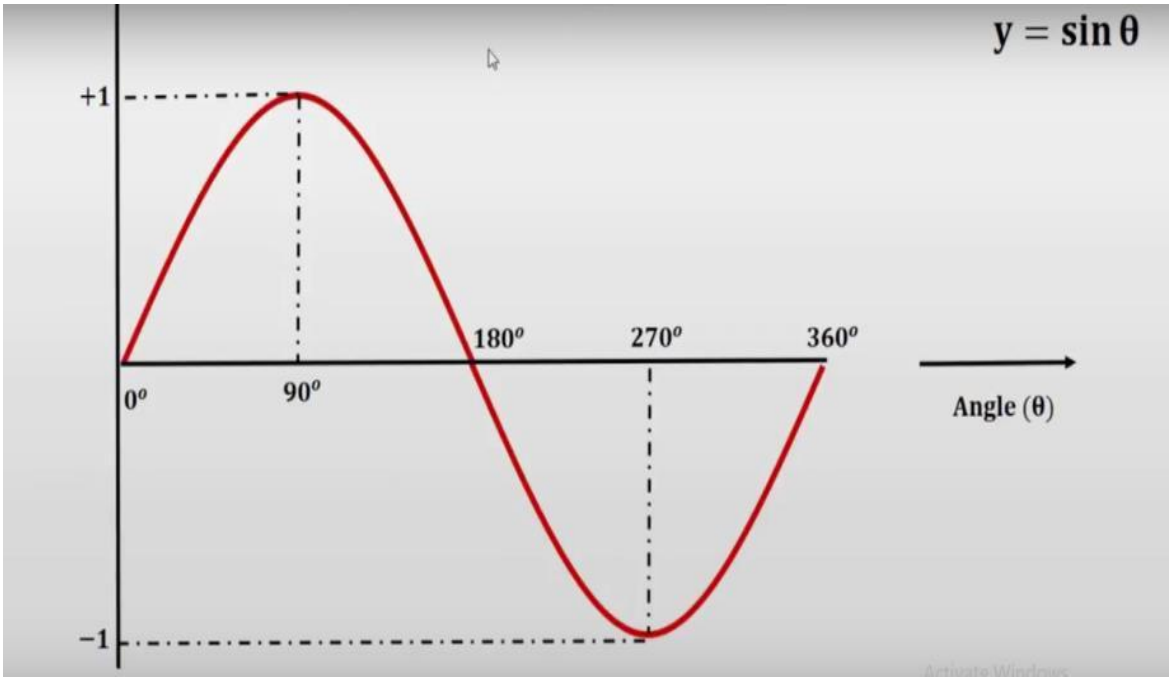
It is half the range of excursion of the cycle with the given period; in a symmetrical oscillation, the amplitude corresponds to the distance between the mesor and the peak (or trough) of the wave.

1.2.4 Phase (or Phase angle):

It is the relative angular displacement between the oscillation and a reference angle.

 It is the position of a point in time (instant) on a waveform cycle.

- In given figure, curve is sinusoidal its highest value is +1 and lowest value is -1.



- **Acrophase :**It is the phase of the maximal value assumed by the curve.
- **Orthophase :** It is the phase of the maximum in the case of a composite model.
- **Bathyphase :** It is the phase of the minimal value assumed by the curve.

1.2.5 Phase difference:

The difference between two sinusoidally varying quantities that have the same frequency, measured either as an angle or a time.

1.2.6 Waveform:

- It refers simply to the shape of the wave.
- It is a graphical representation of a signal in the form of a wave.
- It can be both sinusoidal as well as square shaped.

- Waveforms follow a mathematical function that defines how they are represented and allowed to be interpreted by the reader.

For example,

- A sinusoidal waveform follows a trigonometric function.
- The square waveform follows a harmonic function.

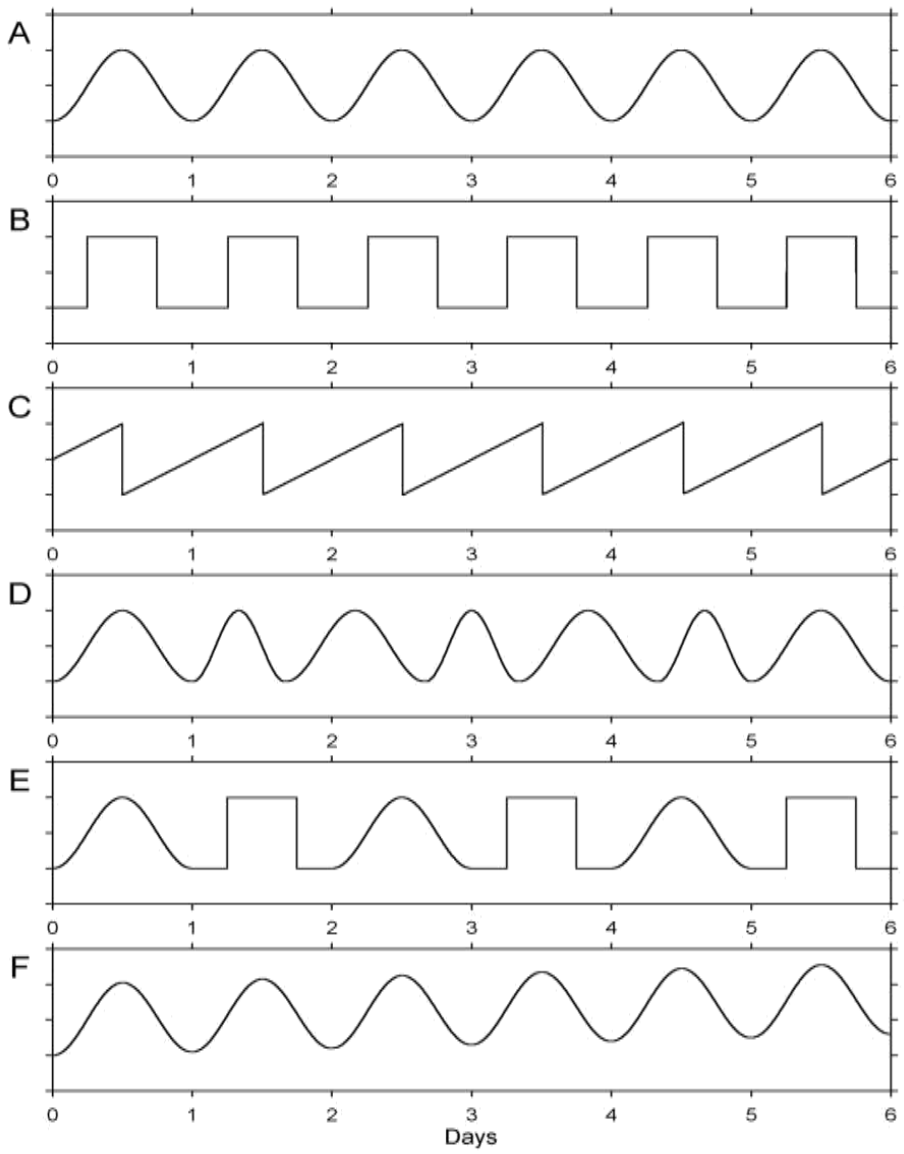
Types of wave forms:

In the following figure,

- Waves A and B have the same mesor, same magnitude, same phase (on average), and same period.

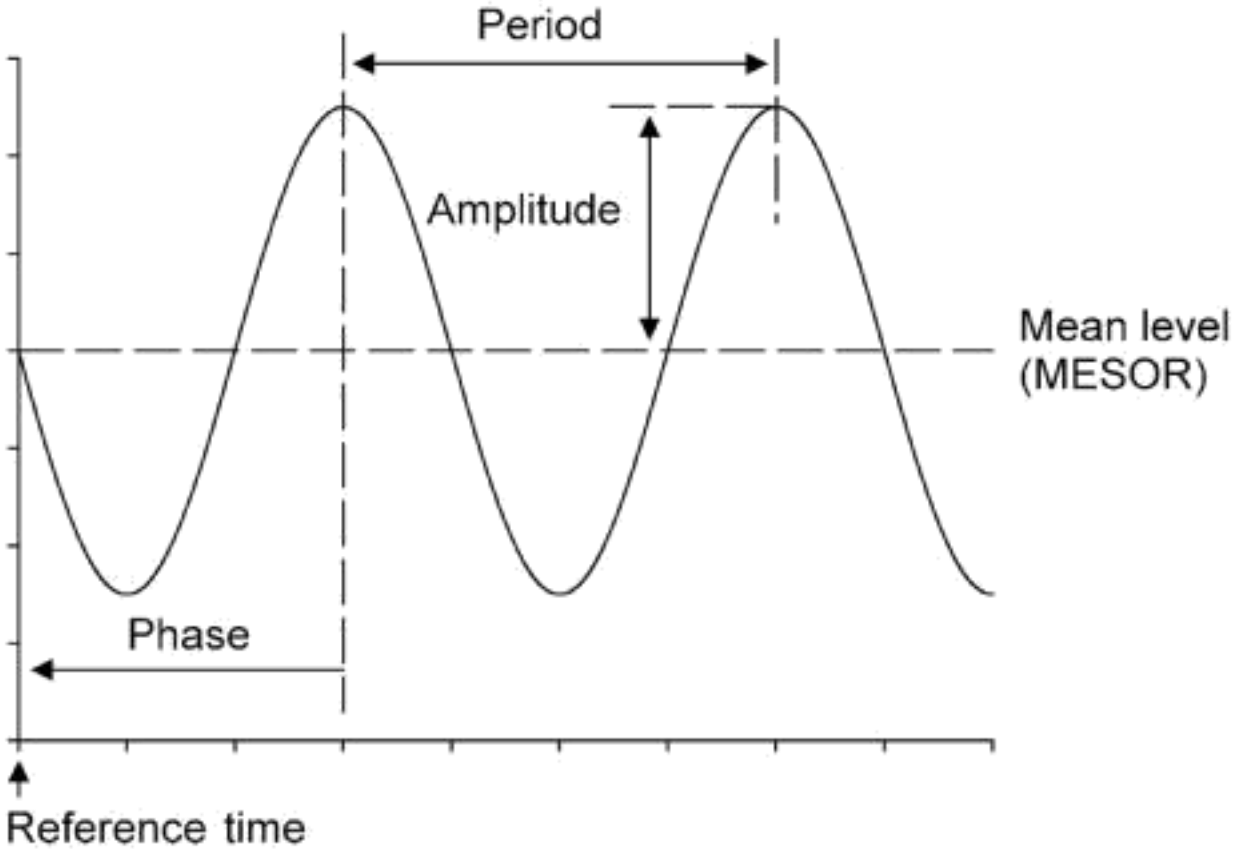
- Yet, they are clearly different from each other.

- Their waveforms are different:
  A is sinusoidal and B is square, Wave D have unstable period (D) and wave E have unstable waveform. There is linear trend in F.



- Difference between A and F is that, in linear trend one can observe the mesor drift which is not the case in sinusoidal waveform.

Following figure diagrammatically describes the terms explained above.



1.3 Types of sampling

There are essentially three types of sampling of time series:

1.3.1 Transverse (cross-sectional) sampling

- It is sampling of many individuals, once per individual.
- To return to our example, we might choose to measure cholesterol levels in daily walkers across two age groups, over 40 and under 40, and compare these to cholesterol levels among non-walkers in the same age groups. We might even create subgroups for gender.
- However, we would not consider past or future cholesterol levels, for these would fall outside the frame. We would look only at cholesterol levels at one point in time.
- The benefit of a cross-sectional study design is that it allows researchers to compare many different variables at the same time. We could, for example, look at age, gender, income and educational level in relation to walking and cholesterol levels, with little or no additional cost.

1.3.2 Longitudinal sampling

- It is sampling conducted continuously over many cycles, preferably at regular intervals.
- In our example, we might choose to look at the change in cholesterol levels among women over 40 who walk daily for a period of 20 years.
- The longitudinal study design would account for cholesterol levels at the onset of a walking regime and as the walkingbehaviour continued over time.
- Therefore, a longitudinal study is more likely to suggest cause-and-effect relationships than a cross-sectional study by virtue of its scope.
- The benefit of a longitudinal study is that researchers are able to detect developments or changes in the characteristics of the target population at both the group and the individual level.
- The key here is that longitudinal studies extend beyond a single moment in time. As a result, they can establish sequences of events.

1.3.3 Hybrid sampling

- Hybrid designs are a combination of longitudinal and transverse sampling.
- In hybrid sampling, we have many subgroups and we conducted sampling continuously over many cycles.
- Example: Panel election data.

We make subgroups of voter according to their economical condition, region wise and gender. Then we make survey for every year each new survey can be compared to previous survey responses. This allows us to gain a better understanding of changes in voting intentions and behavior.

Comparison between Cross sectional and Longitudinal Sampling

Cross Sectional	**Longitudinal Sampling**
One point in time.	Several point in time.
Different Sample.	Same Sample.
Snapshot of given point in time, change at social level.	Change at individual level.

2. Motivation

- Your circadian rhythm works best when you have regular sleep habits, like going to bed at night and waking up in the morning around the same times from day to day (including weekends). When things get in the way, like jet lag, daylight savings time, or a compelling sporting event on TV that keeps you up into the wee hours of the morning, you can disrupt your circadian rhythm, which makes you feel out of sorts and can make it harder to pay attention.

- Interestingly, your circadian rhythm will likely change as you get older. And you may not have the same sleep/wake cycle as your partner, child or parents. But the more you pay attention to your body and notice feelings of alertness and drowsiness, and the more time you spend developing good sleep hygiene habits, the better your slumber will be and the better you'll feel.

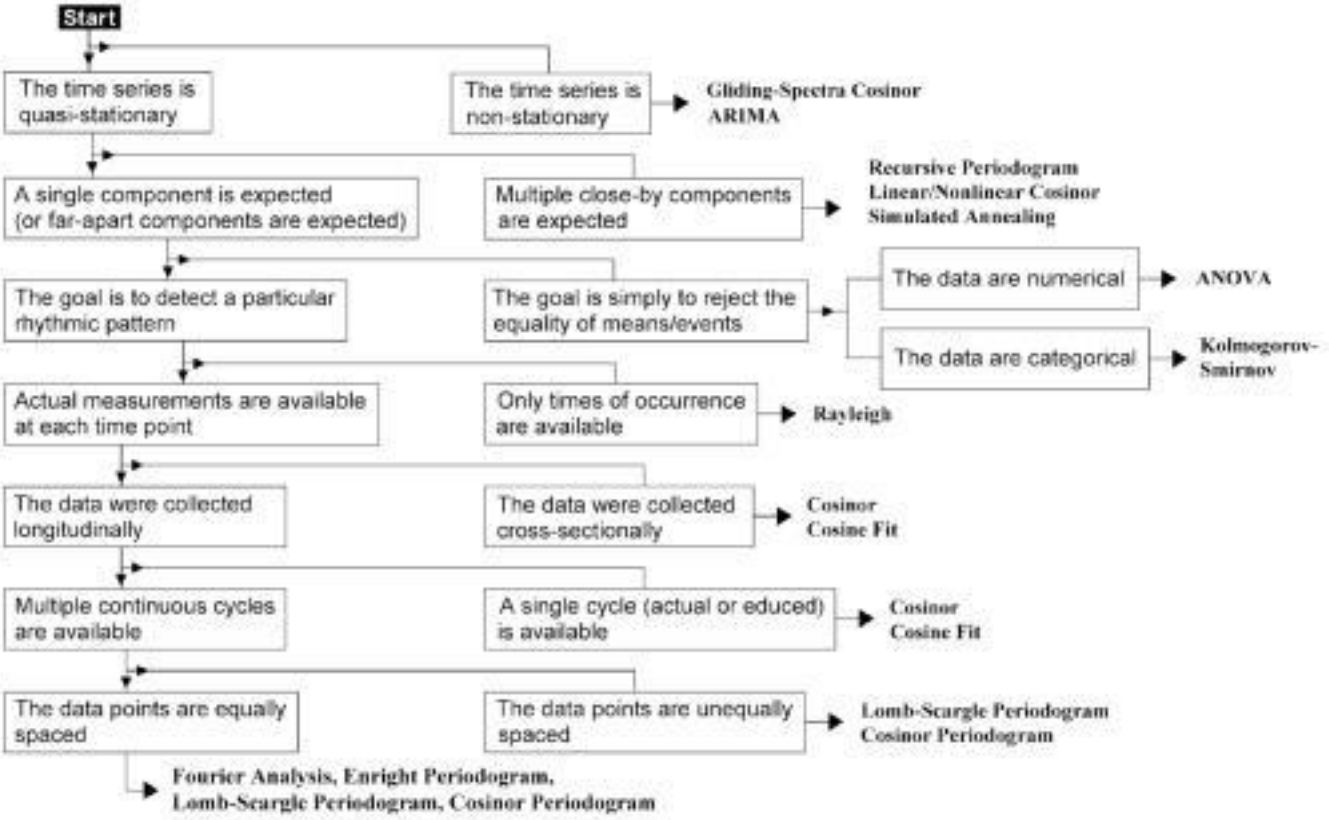
- In a strict sense, circadian rhythms are endogenously generated, although they can be modulated by external cues such as sunlight and temperature.

- Circadian rhythms are important in determining the sleeping and feeding patterns of all animals, including human beings.

- There are clear patterns of brain wave activity, hormone production, cell regeneration and other biological activities linked to this daily cycle.

- Mood disorders are serious diseases that affect a large portion of the population. Nearly all people suffering from mood disorders have significant disruptions in circadian rhythms and the sleep/wake cycle.

3. Selection of procedures for the detection



4. Data description

4.1 About the data

4.1.1 Temperature data set

In this data set, two types of temperature are measured. They are palm and oral temperature. Measurements were taken for 59 days from 1st January to 15th March. For each of the day, measurements were taken at 3 different times (12:00noon,6:00PM, 11:00PM) at irregular intervals.

4.1.2 Blood pressure data set

Two types of Blood pressures (Systolic and Diastolic) were measured on 3 different patients for 59 days,from 5th December to 2nd February. For each of the day, measurements were taken at 7 different times (6:00AM,8:00AM,12:00PM,3:00PM,6:00PM,9:00PM, 11:00PM) at irregular intervals.

4.2 Modifications Made

In above all data sets, we have converted the time variables as continuous hours for Lomb Scargle Periodogram.

Example: If first measurement was taken at 6 AM then it is taken as 6 and so on. Further, for each consecutive day, we have added the times to convert into hours.

Actual Time Variable	Modified time Variable
6:00 AM	6
8:00 AM	8
12:00 PM	12
3:00 PM	15
6:00 AM	18
9:00 PM	21
11:00 PM	23
6:00 AM	30
8:00 AM	32

For Kruskalwallis test and cosinorperiodogram, we have taken time points as given.

5.Techniques applied

5.1 Significance of effect of time on blood pressure using Kruskal - Wallis test

- Null hypothesis H_0: median of blood pressure at time 6:00 AM, 8:00 AM, 12:00 PM, 3:00 PM, 6:00 PM, 9:00 PM and11:00 PM is same i.e. There is no significant effect of time on blood pressure
  Vs
  Alternative hypothesis H_1: At least one median is different i.e. There is significant effect of time on blood pressure
- Test statistic is

$$H = \left[\frac{12}{n(n+1)} \sum_{j=1}^{c} \frac{T_j^2}{n_j} \right] - 3(n+1)$$

where

n = sum of sample sizes for all samples,

c = number of samples,

T_j = sum of ranks in the j^{th} sample,

n_j = size of the j^{th} sample.

- The z-value for each group is calculated as:

$$Z_j = \frac{\bar{R}_j - \bar{R}}{\sqrt{\frac{(N+1)\left(\frac{N}{n_j} - 1\right)}{12}}}$$

The value of z_j indicates how the mean rank, R_j, for group j differs from the mean rank, R, for all N observations.

5.2Lomb Scargleperiodogram

- Technique better suited for –Unequally sampled time series and data sets with missing values.

- Benefit of providing easy way to calculate ‘false alarm probabilities’.

- False Alarm Probablity- Typically, the data analysed are noisy. As a result, spurious peaks arise in periodograms of the data, not because of any periodicity in the observed system, but because of the way that the noisy signal has been sampled. the probability of this nouisy signals are false alarm probability. These spurious peaks can be surprisingly large. It is essential therefore to have reliable tests by which to determine the significance of periodogram peaks.

USES

- Better detection efficiency and accuracy in presence of noise.
- Avoids possible bias & erroneous results arising from replacement of missing data by interpolation.

Lomb Scargle statistic

- For a data set with N values (x_i ; i =1 to N) collected at time points t_i , the PN statistic is computed as :

$$PN=\frac{1}{2\sigma^2}\left\{\frac{\left[\sum_{i=1}^{N}(x_i-M)\cos(\frac{2\pi}{P}(t_i-\delta))\right]^2}{\sum_{i=1}^{N}\cos^2(\frac{2\pi}{P}(t_i-\delta))}+\frac{\left[\sum_{i=1}^{N}(x_i-M)\sin(\frac{2\pi}{P}(t_i-\delta))\right]^2}{\sum_{i=1}^{N}\sin^2(\frac{2\pi}{P}(t_i-\delta))}\right\},$$

where

M is the mean of all N values

σ^2 is the variance of all N values

P is the period being tested

δ is a term used to adjust the times of unequally spaced data

$$\delta=\frac{1}{\frac{4\pi}{P}}\tan^{-1}\left[\frac{\sum_{i=1}^{N}\sin(\frac{4\pi}{P}t_i)}{\sum_{i=1}^{N}\cos(\frac{4\pi}{P}t_i)}\right]$$

- The significance level of the periodogram,

$$p(PN)=1-(1-e^{-PNmax})^N$$

wherePN_{max} is the largest PN in the periodogram.

Interpretations

- In thalsp (Lomb scagrleperiodogram)the period corresponding to highest normalized power is the true estimated period.
- Visual inspection of periodogram suggest the presence of strong rhythemicity with period of 24 hours.
- The dashed line in the periodogram indicate the 0.01 l.o.s. it is calculated by formula *p(PN)*
- If the Pval<0.01. Hence, rhythmicity with period of 24 hours is significant.
- Also in the periodogram the peak must be exceed to **Λ** to reach significance alpha calculated by formula

$$\Lambda = -\ln\left[1 - (1-\alpha)^{\frac{1}{M}}\right]$$

whereM is sample size.

SCARGLE'S SIGNIFICANCE TEST

- If the data are Gaussian pure noise, the periodogram power $Z = PX(\omega)$ at any given frequency ω of the sampled signal *Xk*is exponentiallydistributed with probability density function defined by

$p(z)\,dz = \Pr[z < Z < z + dz]$

$$= \frac{1}{\sigma_X^2} e^{-z/\sigma_X^2}\, dz.$$

- The CDF is thus given by,

$$P_Z(z) = \Pr[Z < z]$$
$$P(Z) = 1 - e^{-z/\sigma 2}$$

- We are interested in the probability that the periodogram power atthe given frequency is greater than a specified threshold *z*. This isgiven by, $\Pr[Z > z] = 1 - P(z) = e^{-z/\sigma 2}$

- To calculate the probability that all the sampledperiodogram powers are less than some specified threshold powerz, define a new random variable

$$Z_{\max} = \text{maximum}\{Z_1, Z_2, \ldots, Z_{N_i}\}.$$

- Suppose now that we evaluate the periodogram at frequencies$\{\omega\mu{:}\mu = 1, 2, \ldots, Ni\}$. Denote the periodogram powers at thesefrequencies by $Z\mu = PX(\omega\mu)$.

- The probability that any given power $Z\mu$in this set falls below thethreshold is

$$\Pr[Z_\mu < z] = 1 - e^{-z/\sigma_X^2}.$$

- Since the $Z\mu$are independent, the probability that they all fall below the threshold z is given by

$$\begin{aligned}\Pr[Z_1 < z \text{ and } Z_2 < z \text{ and} \ldots \text{and } Z_{N_i} < z] \\ = \Pr[Z_1 < z]\Pr[Z_2 < z] \ldots \Pr[Z_{N_i} < z] \\ = \left[1 - e^{-z/\sigma_X^2}\right]^{N_i}.\end{aligned}$$

- The probability that not all the powers $Z\mu$are less than the thresholdz, that is, the probability that at least one of the powers $Z\mu$is abovethe threshold z, is then

$$\Pr[Z_{\max} > z] = 1 - \left[1 - e^{-z/\sigma_X^2}\right]^{N_i}. \quad \text{---------(A)}$$

- This is the function that Scargle proposes as a false alarm probability function that we regard as an acceptable level of risk for the false detectionof real deterministic signals; then solve (A) for z to get a referencepower threshold level z_A:

$$z_A = -\sigma_X^2 \ln\left[1 - (1 - p_A)^{1/N_i}\right].$$

- If we claim a detection whenever the power level at one of thefrequencies $\{\omega_\mu : \mu = 1, 2, \ldots, N_i\}$ exceeds the reference level z_A,the probability that we will be wrong is given by p_A.

5.3 Cosinor Periodogram

- The cosinor method was developed by Franz Halberg and colleagues in the early 1960s.

- It is appropriate not only for non-equidistant data but also for serially independent data.

- The method is based on the reasoning that, since circadian rhythms can be thought of as smooth rhythms with added noise, a model consisting of cosine curves with known periods (24 h alone or with added harmonic terms) can be fitted by least squares to the data as an estimate of the pattern of the smooth rhythm.

- In the special case of a single cosinusoidal curve, the value of each point (x_i) is a function of the average value of the variable under investigation, estimated as the MESOR(M), the amplitude of the oscillation (A), and the acrophase (φ).

- Since the period (P) is assumed known (or approximately known), the fit can be obtained by linear regression --- the trigonometric angles (θ_i) corresponding to the sampling times t_i (at which the data points xi were collected) being computed as

  $\theta_i=(2\pi\ t_i)/P$.

- The model can be written according to the equation:

$$x_i=M+A\cos(\theta_i+\varphi)+e_i$$

  wheree_i is an error term assumed to be independent and normally distributed with mean zero and fixed unknown variance σ^2.

- If this equation is incorporated into a model of linear regression by the method of least squares, a system of three equations with three unknowns can be derived and solved in algebraic form to provide the parameters of the cosine wave that best fits the time series.
- When the period is known, the model can be rewritten as

  $$\mathbf{x_i = M + \beta X_i + \gamma Z_i + e_i}$$

  where $\beta = A\cos\varphi$; $\gamma = -A\sin\varphi$; $X_i = \cos(2\pi t_i/P)$; $Z_i = \sin(2\pi t_i/\tau)$

  The model is linear in its parameters M, β and γ.

- Solving the system of equations thus obtained for the parameters M, β and γ can be done with matrix algebra, or using a linear model function. The following demonstrates the method for solving by matrix.

System of equations:

$$\Sigma y_i = MN + \beta \Sigma \cos\omega t_i + \gamma \Sigma \sin\omega t_i$$

$$\Sigma y_i \cos\omega t_i = M \Sigma \cos\omega t_i + \beta \Sigma \cos^2\omega t_i + \gamma \Sigma \sin\omega t_i \cos\omega t_i$$

$$\Sigma y_i \sin\omega t_i = M \Sigma \sin\omega t_i + \beta \Sigma \sin\omega t_i \cos\omega t_i + \gamma \Sigma \sin^2\omega t_i$$

System of equations, where $\omega=2\pi/P$

Matrix form:

$$\begin{pmatrix} \Sigma y_i \\ \Sigma y_i \cos\omega t_i \\ \Sigma y_i \sin\omega t_i \end{pmatrix} = \begin{pmatrix} N & \Sigma \cos\omega t_i & \Sigma \sin\omega t_i \\ \Sigma \cos\omega t_i & \Sigma \cos^2\omega t_i & \Sigma \sin\omega t_i \cos\omega t_i \\ \Sigma \sin\omega t_i & \Sigma \sin\omega t_i \cos\omega t_i & \Sigma \sin^2\omega t_i \end{pmatrix} \begin{pmatrix} M \\ \beta \\ \gamma \end{pmatrix}$$

Solving for M, β, γ :

$$\begin{pmatrix} M \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} N & \Sigma \cos\omega t_i & \Sigma \sin\omega t_i \\ \Sigma \cos\omega t_i & \Sigma \cos^2\omega t_i & \Sigma \sin\omega t_i \cos\omega t_i \\ \Sigma \sin\omega t_i & \Sigma \sin\omega t_i \cos\omega t_i & \Sigma \sin^2\omega t_i \end{pmatrix}^{-1} \begin{pmatrix} \Sigma y_i \\ \Sigma y_i \cos\omega t_i \\ \Sigma y_i \sin\omega t_i \end{pmatrix}$$

- If the fitted wave has amplitude statistically greater than zero, then the time series can be inferred to exhibit 24-h rhythmicity.

- The cosinor computations are greatly simplified if the data points are equidistant and cover an integer number of cycles. In this case, the MESOR (M) is simply the arithmetic mean of all data points [M = (Σxi)/N, for i =1 to N] and A and φ are computed as:

$$\widehat{A} = \sqrt{\widehat{\beta}^2 + \widehat{\gamma}^2} \text{ and}$$

$$\begin{aligned} \varphi &= -\tan^{-1}\left|\frac{\widehat{\gamma}}{\widehat{\beta}}\right| && \text{if } \widehat{\gamma} > 0 \text{ and } \widehat{\beta} \geq 0 \\ &= -\pi + \tan^{-1}\left|\frac{\widehat{\gamma}}{\widehat{\beta}}\right| && \text{if } \widehat{\gamma} > 0 \text{ and } \widehat{\beta} < 0 \\ &= -\pi - \tan^{-1}\left|\frac{\widehat{\gamma}}{\widehat{\beta}}\right| && \text{if } \widehat{\gamma} < 0 \text{ and } \widehat{\beta} \leq 0 \\ &= -2\pi + \tan^{-1}\left|\frac{\widehat{\gamma}}{\widehat{\beta}}\right| && \text{if } \widehat{\gamma} \leq 0 \text{ and } \widehat{\beta} > 0 \end{aligned}$$

where

$$\widehat{\beta} = \frac{2}{N}\sum_{i=1}^{N} \cos(\theta_i) x_i \quad \text{and} \quad \widehat{\gamma} = \frac{2}{N}\sum_{i=1}^{N} \sin(\theta_i) x_i.$$

- The probability that the amplitude (A) is significantly different from zero can be calculated by the F distribution with 2 and N – 3 degrees of freedom because

$$\frac{N(\widehat{\beta}^2+\widehat{\gamma}^2)}{4\widehat{S}^2} \approx F_{2,N-3},$$

where S^2 is the residual variance computed as:

$$\widehat{S}^2=\frac{\sum_{i=1}^{N}\left\{x_i-[\widehat{M}+\widehat{A}(\cos\varphi\cdot\cos\theta_i-\sin\varphi\cdot\sin\theta_i)]\right\}^2}{(N-3)}.$$

6. Analysis

6.1 Kruskal- Wallis Test

6.1.1 For Mithilesh-Systolic BP

Kruskal-Wallis Test: sys bp versus time

Time	N	Median	Ave Rank	Z
6:00 AM	60	127.0	274.5	4.41
8:00 AM	60	126.0	237.0	1.83
12:00 PM	60	125.0	217.2	0.46
3:00 PM	60	121.5	149.2	-4.23
6:00 PM	60	122.0	154.8	-3.84
9:00 PM	60	123.0	153.7	-3.91
11:00 PM	60	129.5	287.2	5.28
Overall	420		210.5	

```
H = 84.70    DF = 6 P = 0.000
H = 85.10    DF = 6 P = 0.000  (adjusted for ties)
```

- P-value is < 0.05, Therefore we reject H_0 i.e. There is significant effect of time on blood pressure.
- The z value at time 12:00 PM is 0.46, the smallest absolute z-value.This size indicates that the mean rank at time 12.00PM differed least from the mean rank for all observations i.e., systolic blood pressure at 12:00PM is close to averagesystolic blood pressure of all time points.

- The mean rank at time 3.00PM is lower than the mean rank for all observations i.e. systolic blood pressure at 3:00 PM is significantly low than the averagesystolic blood pressure of all time points.

- The mean rank at time 11.00PM is higher than the mean rank for all observations, as the z-value is positive (5.28) i.e. systolic blood pressure at 11:00 PM is significantly high than the average systolicblood pressure of all time points.

6.1.2 ForMithilesh-Diastolic BP

Kruskal-Wallis Test: diabp versus time

Time	N	Median	Ave Rank	Z
6:00 AM	60	78.00	226.5	1.10
8:00 AM	60	73.50	169.0	-2.86
12:00 PM	60	79.00	212.8	0.16
3:00 PM	60	75.00	178.6	-2.20
6:00 PM	60	80.00	249.9	2.71
9:00 PM	60	78.00	221.8	0.78
11:00 PM	60	76.00	214.8	0.30
Overall	420		210.5	

H = 19.13 DF = 6 P = 0.004
H = 19.19 DF = 6 P = 0.004 (adjusted for ties)

- P-value is < 0.05, Therefore we reject H_0 i.e. There is significant effect of time on blood pressure.
- The z value at time 12PM is 0.16, the smallest absolute z-value.This size indicates that the mean rank at time 12.00PM differed least from the mean rank for all observations i.e. diastolic blood pressure at 12:00PM is close to average diastolic blood pressure of all time points.

- The mean rank at time8.00AM is lower than the mean rank for all observations i.e. diastolic blood pressure at 8:00 AM is significantly low than the averagediastolic blood pressure of all time points.
- The mean rank at time 6.00PM is higher than the mean rank for all observations, as the z-value is positive (2.71) i.e. diastolic blood pressure at 6:00 PM is significantly high than the average systolic blood pressure of all time points.

6.1.3 ForApoorv-Systolic BP

Kruskal-Wallis Test: sys versus time

Time	N	Median	Ave Rank	Z
6:00 AM	60	131.0	240.8	2.09
8:00 AM	60	128.0	229.5	1.31
12:00 PM	60	130.0	240.7	2.08
3:00 PM	60	125.0	155.6	-3.78
6:00 PM	60	124.0	141.6	-4.75
9:00 PM	60	129.0	229.4	1.30
11:00 PM	60	129.0	235.8	1.74
Overall	420		210.5	

H = 44.60 DF = 6 P = 0.000
H = 44.79 DF = 6 P = 0.000 (adjusted for ties)

- P-value is < 0.05, Therefore we reject H_0 i.e. There is significant effect of time on blood pressure.
- The z value at 9PM is 1.30which is the smallest absolute z-value.This indicates that the mean rank at time 9.00PM differed least from the mean rank for all observations i.e. systolic blood pressure at 9:00PM is close to average systolic blood pressure of all time points.

- The mean rank at time 6.00PM is lower than the mean rank for all observations i.e. systolic blood pressure at 6:00 PM is significantly low than the averagesystolic blood pressure of all time points.

- The mean rank at time 6.00AM is higher than the mean rank for all observations, as the z-value is positive (2.09) i.e. systolic blood pressure at 6:00 AM is significantly high than the averagesystolic blood pressure of all time points.

6.1.4 ForApoorv-Diastolic BP

Kruskal-Wallis Test: dia versus time

Time	N	Median	Ave Rank	Z

6:00 AM	60	81.00	241.1	2.11
8:00 AM	60	78.00	198.0	-0.86
12:00 PM	60	79.00	196.5	-0.97
3:00 PM	60	78.00	190.1	-1.41
6:00 PM	60	76.00	169.7	-2.81
9:00 PM	60	78.50	198.1	-0.86
11:00 PM	60	83.00	280.1	4.80
Overall	420		210.5	

H = 34.07 DF = 6 P = 0.000
H = 34.17 DF = 6 P = 0.000 (adjusted for ties)

- P-value is < 0.05, Therefore we reject H_0 i.e. There is significant effect of time on blood pressure.
- The z value at 9PM and 8AMis 0.86 which is the smallest absolute z-value.This indicates that the mean rank at time 9.00PM and 8.00AM differed least from the mean rank for all observations. BP at 9PM and 8AM are nearly same.

- The mean rank at time 6.00PM is lower than the mean rank for all observations. Therefore blood pressure is low at 6PM i.e. diastolic blood pressure at 6:00 PM is significantly low than the averagediastolic blood pressure of all time points.

- The mean rank at time 11PM is higher than the mean rank for all observations, as the z-value is positive (4.80). Therefore blood pressure is high at 11PM. i.e. diastolic blood pressure at 11:00 PM is significantly high than the averagediastolic blood pressure of all time points.

6.1.5 ForAkash-Systolic BP

Kruskal-Wallis Test: sys versus time

time	N	Median	Ave Rank	Z
6:00 AM	60	130.0	221.8	0.78
8:00 AM	60	125.0	157.8	-3.63
12:00 PM	60	125.5	161.5	-3.38

3:00 PM	**60**	**128.5**	**195.0**	**-1.07**
6:00 PM	60	131.5	237.3	1.84
9:00 PM	60	130.0	225.8	1.06
11:00 PM	60	135.0	274.4	4.40
Overall	420		210.5	

H = 43.08 DF = 6 P = 0.000
H = 43.18 DF = 6 P = 0.000 (adjusted for ties)

- P-value is < 0.05, Therefore we reject H_0 i.e. There is significant effect of time on blood pressure.
- The z value at 6AM is 0.78 which is the smallest absolute z-value.This indicates that the mean rank at time 6.00AM differed least from the mean rank for all observations i.e. systolic blood pressure at 6:00AM is close to average systolic blood pressure of all time points.

- The mean rank at time 8.00AM is lower than the mean rank for all observations. Therefore blood pressure is low at 8AM i.e. systolic blood pressure at 8:00 AM is significantly low than the averagediastolic blood pressure of all time points.

- The mean rank at time 11.00PM is higher than the mean rank for all observations, as the z-value is positive (4.40) i.e. systolic blood pressure at 11:00 PM is significantly high than the averagesystolic blood pressure of all time points.

6.1.6 ForAkash-Diastolic BP

Kruskal-Wallis Test: dia versus time

Time	**N**	**Median**	**Ave Rank**	**Z**
6:00 AM	60	78.50	195.4	-1.04
8:00 AM	60	81.50	216.0	0.38
12:00 PM	60	80.00	213.2	0.19
3:00 PM	60	78.00	187.6	-1.58
6:00 PM	60	82.00	216.2	0.40
9:00 PM	60	80.00	209.6	-0.06
11:00 PM	60	82.50	235.4	1.71
Overall	420		210.5	

H = 5.87 DF = 6 P = 0.438
H = 5.89 DF = 6 P = 0.436 (adjusted for ties)

- Here p-value is greater than 0.05, So we accept H_0 i.e. there is no significant effect of time on blood pressure.

6.2 Lomb ScarglePeriodogram

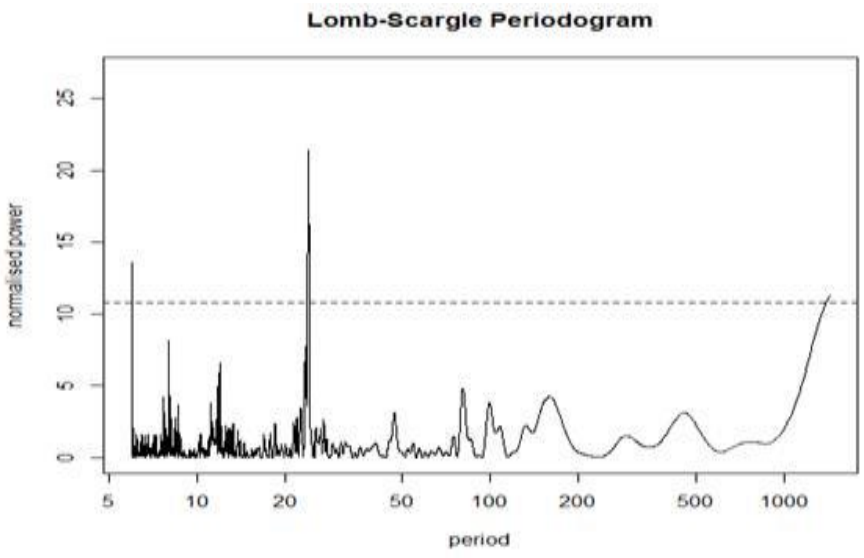
6.2.1 For Mithilesh'sBlood Pressure

- In the data set, number of variables are 3 (Time, Systolic ,Diastolic),number of data points are 406.
- Here we have used Type-"**period"** for constructing periodogram.
- The Oversampling factor is 58 (since the data points are taken at 7 different times for every day and no.of data points are 406.The number of repeating cycles are 406/7 i.e., 58.

- The output of lsp function used for systolic blood pressure :

For Mithilesh'sSystolic BP

```
> summary(LSP)
                      Value
Time                  time1
Data                    Sys
n                       406
Type                 period
Oversampling             58
From                 6.0001
To                     1433
# frequencies         13795
PNmax                21.417
At  period           24.001
At frequency       0.041666
P-value (PNmax) 2.3779e-07
```



Interpretations:

- In the lsp (Lomb scagrle periodogram) the period corresponding to highest normalized power is the true estimated period.

- Visual inspection of periodogram suggest the presence of rhythemicity with the 24 hours.

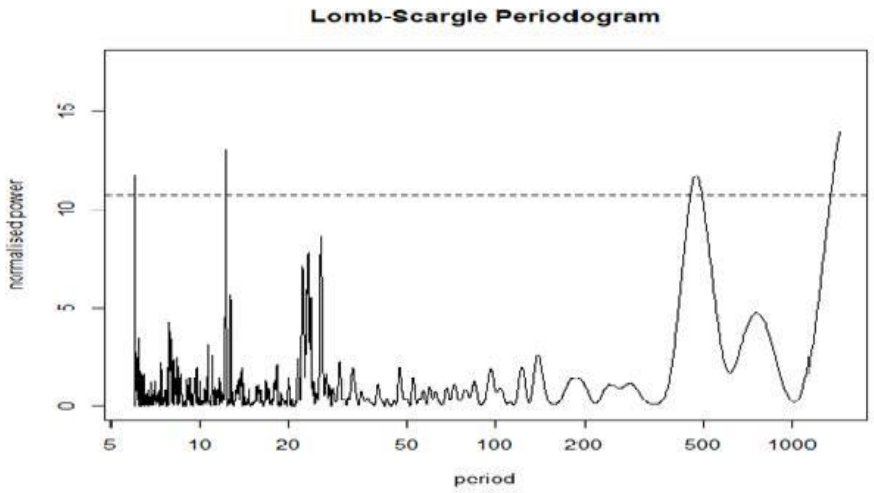
- The dashed line in the periodogram indicate the 0.01 los. it is calculated by formula *p(PN)*.
- PNmax=21.409
- *p(PN)*=2.3779 e-07
- Here Pval<0.01. Hence, rhythmicity with period of 24 hours is significant. Also in the periodogram the peak must be exceed to Λ to reach significance alpha.

$$\Lambda = -\ln\left[1-(1-\alpha)^{\frac{1}{M}}\right]$$

- calculated by formula = 10.60651

For Mithilesh'sDiastolic BP

```
> summary(LSO)
                      Value
Time                  time1
Data                    Dia
n                       406
Type                 period
Oversampling             58
From                 6.0001
To                     1433
# frequencies         13795
PNmax                13.955
At  period             1433
At frequency     0.00069784
P-value (PNmax) 0.00041387
```



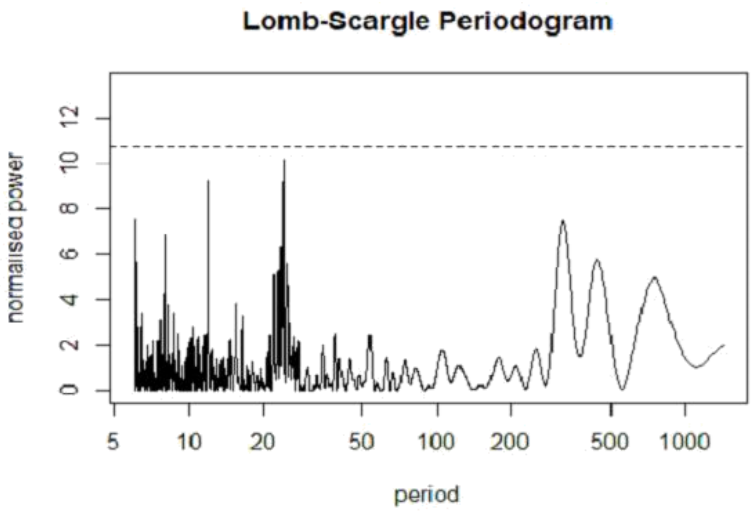
Interpretations:

- Similarly, in the diastolic Blood pressure data, visual inspection of periodogram suggest the presence 1433 hours rhythmicity.
- Here Pval<0.01 hence rhythmicity with period of 1433 hours is significant.

6.2.2 For Apoorva's Blood Pressure

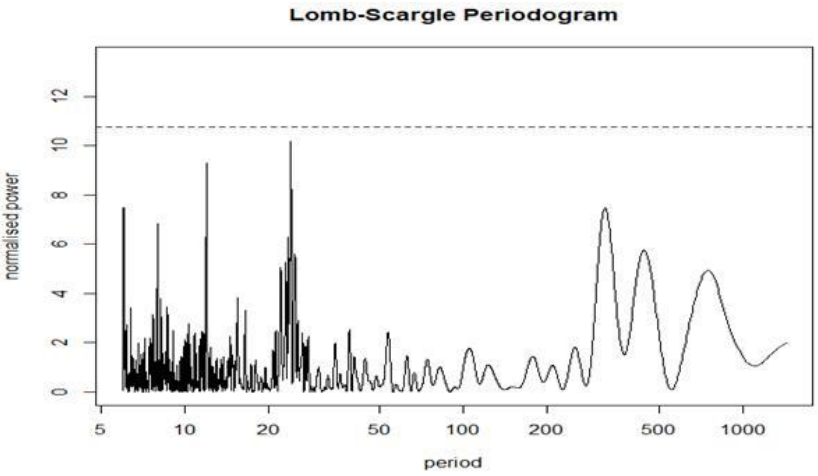
For Apporv'sSystolic BP

```
> summary(LSP)
                    Value
Time                time1
Data                  Sys
n                     403
Type               period
Oversampling           58
From               6.0001
To                   1433
# frequencies       13795
PNmax              10.176
At  period         24.014
At frequency     0.041642
P-value (PNmax) 0.017942
```



For Apporv'sDiastolic BP

```
> summary(LSO)
                    Value
Time                time1
Data                  Dia
n                     403
Type               period
Oversampling           58
From               6.0001
To                   1433
# frequencies       13795
PNmax              10.242
At  period         24.014
At frequency     0.041642
P-value (PNmax) 0.016805
```



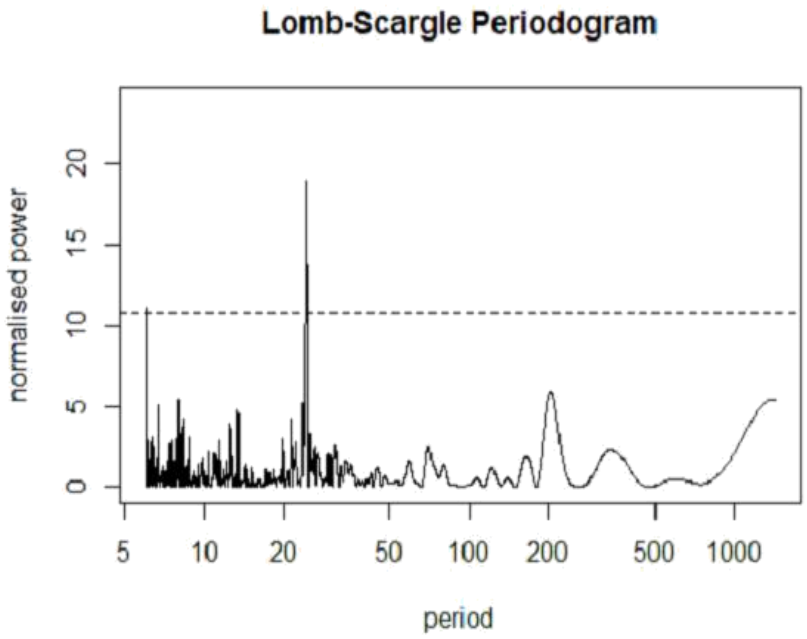
Interpretations:

- Both the Sys and dia. Blood pressure data suggest the presence 24 hours rhythmicity.
- But Here Pval.> 0.01. Hence, rhythmicity with period of 24 hours is not significant in both cases.

6.2.3 For Akash's Blood Pressure

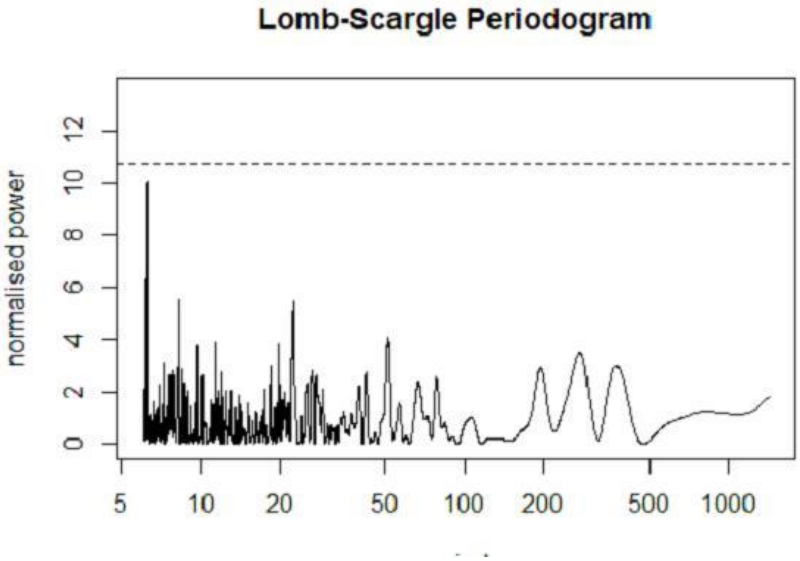
For Akash'sSystolic Blood Pressure

```
> summary(LSP)
                      Value
Time                  time1
Data                    Sys
n                       409
Type                 period
Oversampling             58
From                 6.0001
To                     1433
# frequencies         13795
PNmax                18.944
At  period           24.168
At frequency       0.041377
P-value (PNmax) 2.8185e-06
```



For Akash'sDiastolic Blood Pressure

```
> summary(LSO)
                      Value
Time                  time1
Data                    Dia
n                       409
Type                 period
Oversampling             58
From                 6.0001
To                     1433
# frequencies         13795
PNmax                10.029
At  period           6.2464
At frequency        0.16009
P-value (PNmax) 0.020758
```

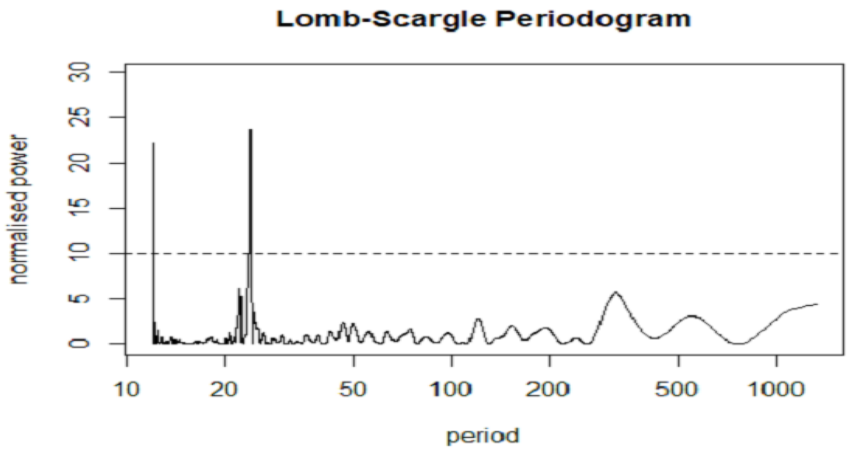


Interpretations:

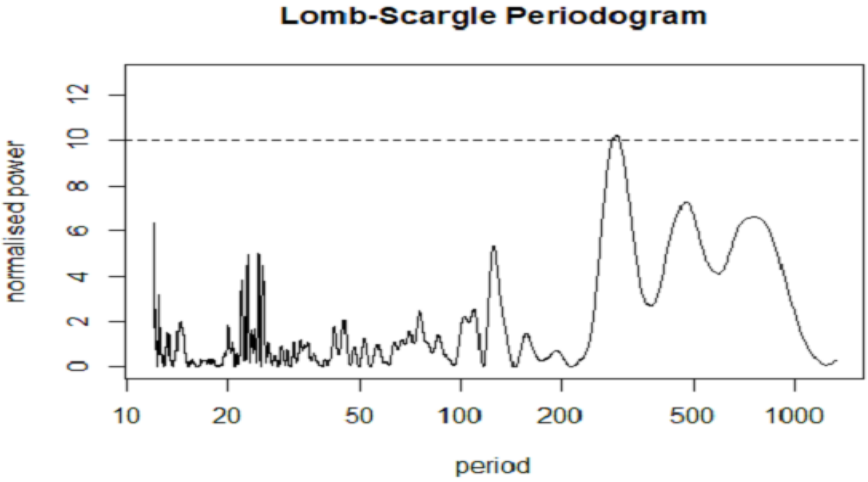
- The Systolic and Diastolic blood pressure data suggest the presence 24 and 6 hours rhythmicity.
- But here p-value> 0.01 for Diastolic. Blood pressure hence rhythmicity with period of 12 hours is not significant.

For Oral Temperature and Palm Termperature

```
> summary(LSO)
                       Value
Time                   time1
Data                    oral
n                        168
Type                  period
Oversampling              56
From                  12.001
To                      1331
# frequencies           6156
PNmax                 23.768
At  period            24.013
At frequency        0.041644
P-value (PNmax) 1.0462e-08
>
```




```
> summary(LSP)
                       Value
Time                   time1
Data                    palm
n                        168
Type                  period
Oversampling              56
From                  12.001
To                      1331
# frequencies           6156
PNmax                 10.247
At  period             292.3
At frequency       0.0034212
P-value (PNmax) 0.0077963
```



Interpretations:

- The oral temperature data suggest the presence 24 hours rhythmicity.
- The Palm temperature data suggest the presence of 292.3 but it is not much significant.

6.3 CosinorPeriodogram

6.3.1 For Mithilesh's Blood pressure

For Mithilesh's Systolic BP

Data consists of 420 readings, but 14 are the missing values. Hence, we have 406 readings with corresponding times at which readings were taken.

Since it is a circadian rhythm, we have taken the period to be 24 hours.

Following is the summary for the cosinor fit to the readings :

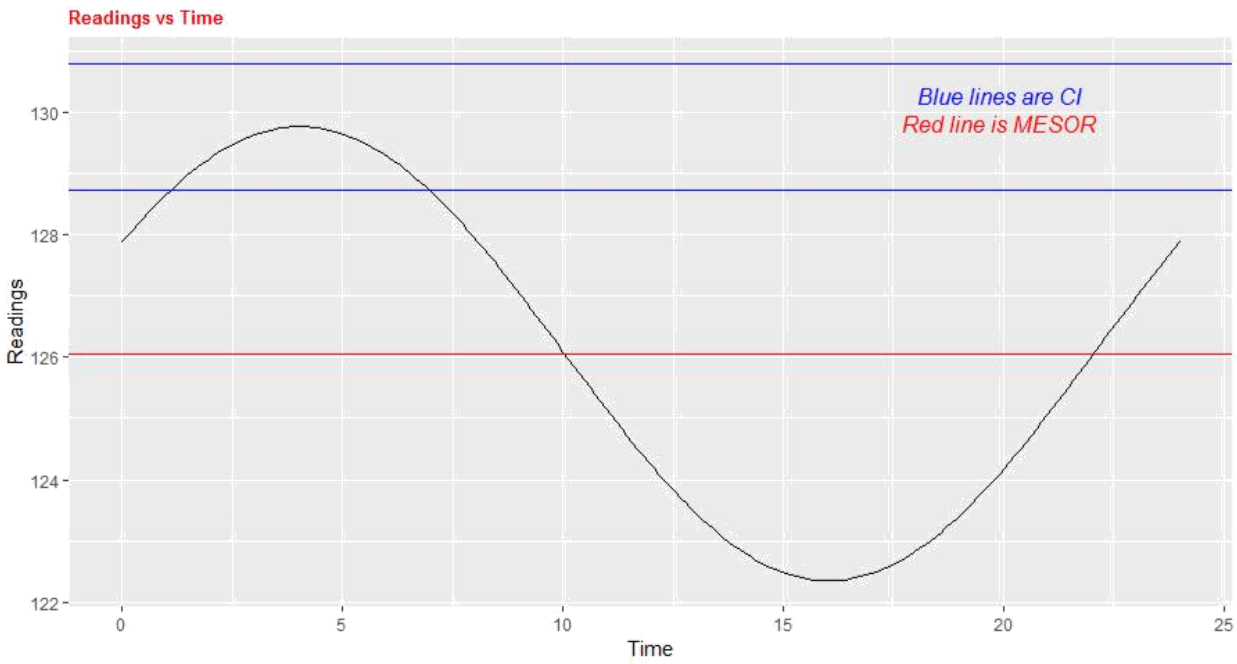
```
Raw model coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept) 126.0611          0.3433 125.3882  126.7341  0e+00
rrr             1.8251          0.5159   0.8141    2.8362  4e-04
sss             3.2369          0.4704   2.3149    4.1589  0e+00

**********************

Transformed coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept) 126.0611          0.3433 125.3882  126.7341       0
amp             3.7160          0.5276   2.6819    4.7501       0
acr             1.0574          0.1230   0.8162    1.2985       0
```

Interpretations:

From the above summary,
MESOR is the intercept = 126.0611
Amplitude = 3.7160
Acrophase = 1.0574 (in radians)
= 60.585 degrees
= 4.04 hours
= 4 hrs 2 mins

For conversions, we have used
1 degree = 180 * 1 radian /3.14 (i.e., 1 rad = 57.2958 degrees)
1 hour =15 degrees
1 degree = 0.07 hour
1 hour = 60 min

According to p-values, we reject the null hypothesis at 5% level of significance. Since, p-values are less than alpha = 0.05 ie., 5% level of significance. Hence, MESOR, amplitude and acrophase are statistically significant from zero.



Also, the 95% confidence interval of the amplitude, shown at the peak, does not overlap the MESOR which implies amplitude of the fitted curve is significantly greater than zero.

For Mithilesh's Diastolic BP

Data consists of 420 readings, but 14 are the missing values. Hence, we have 406 readings with corresponding times at which readings were taken.

Since it is a circadian rhythm, we have taken the period to be 24 hours.

Following is the summary for the cosinor fit to the readings :

```
Raw model coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept)   76.9905         0.3638  76.2774  77.7036 0.0000
rrr            1.1671         0.5466   0.0958   2.2384 0.0327
sss           -0.2325         0.4985  -1.2094   0.7445 0.6410

**********************

Transformed coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept)   76.9905         0.3638  76.2774  77.7036 0.0000
amp            1.1900         0.5231   0.1648   2.2152 0.0229
acr           -0.1966         0.4396  -1.0582   0.6650 0.6547
```

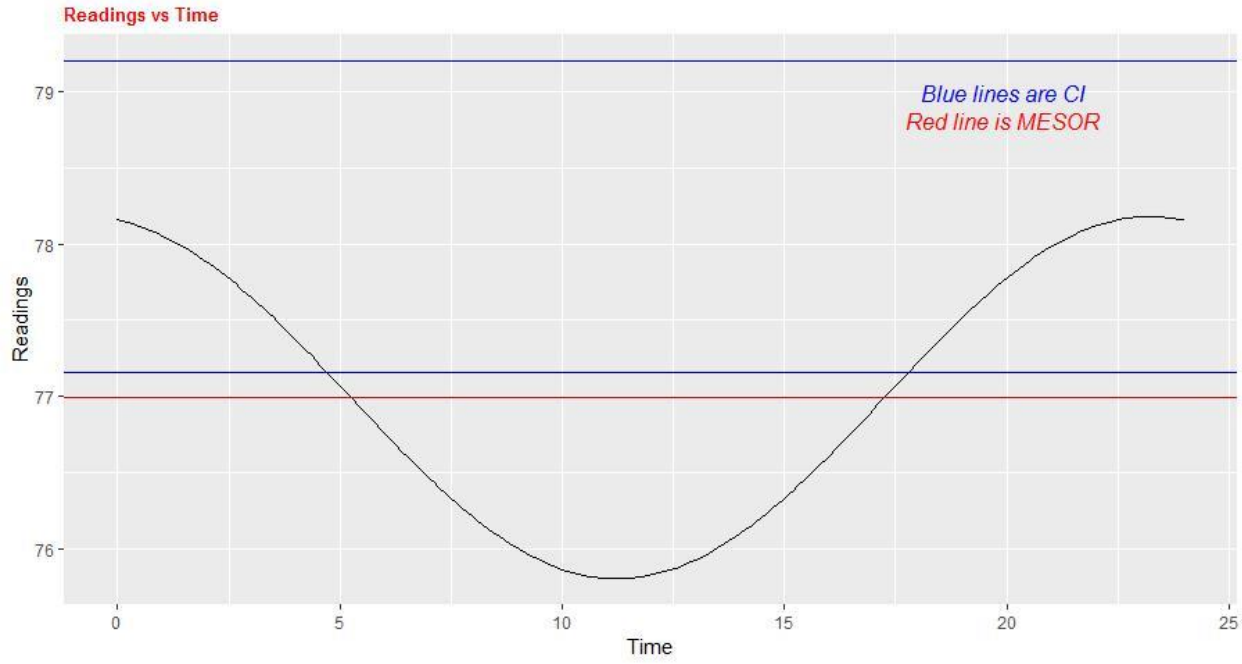
Interpretations:

From the above summary,
MESOR is the intercept = 76.9905

Amplitude = 1.1900
Acrophase = -0.1966 (in radians)
= -11.264 degrees
= -0.75 hours
= -45 mins
= 23 h 15 min

According to p-values, we reject the null hypothesis for MESOR and amplitude at 5% level of significance. Since, p-values are less than alpha = 0.05 ie., 5% level of significance. Hence, MESOR and amplitude are statistically significant from zero.

But, acrophase is insignificant which means peak is at the initial reading so the time corresponding to which peak is occurring is very close to zero (nearly equal to zero).



Also, the 95% confidence interval of the amplitude, shown at the peak, does not overlap the MESOR which implies amplitude of the fitted curve is significantly greater than zero.

6.3.2 For Apoorv's Blood pressure

For Apoorv's Systolic BP

Data consists of 420 readings, but 16 are the missing values. Hence, we have 404 readings with corresponding times at which readings were taken.

Since it is a circadian rhythm, we have taken the period to be 24 hours.

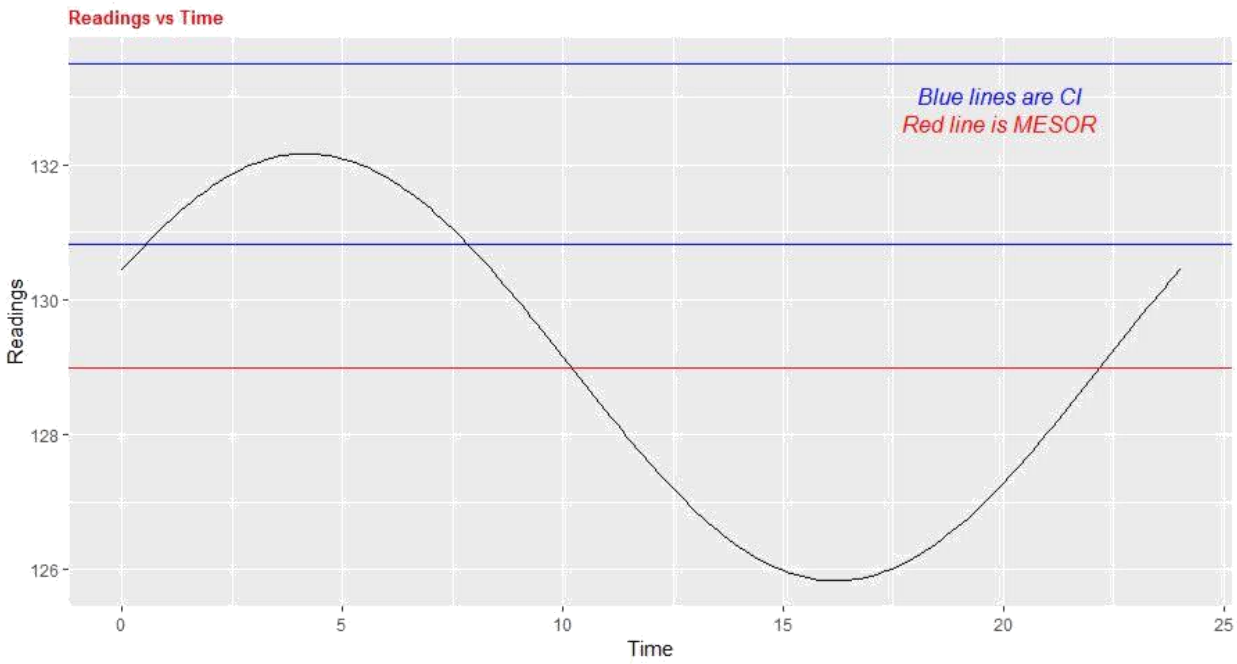
Following is the summary for the cosinor fit to the readings :

```
Raw model coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept) 128.9981          0.4492 128.1177  129.8785 0.0000
rrr            1.4475         0.6753   0.1239    2.7712 0.0321
sss            2.8170         0.6136   1.6144    4.0196 0.0000

**********************

Transformed coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept) 128.9981          0.4492 128.1177  129.8785       0
amp            3.1671         0.6830   1.8284    4.5058       0
acr            1.0961         0.1910   0.7217    1.4705       0
```

Interpretations:

From the above summary,
MESOR is the intercept = 128.9981
Amplitude = 3.1671
Acrophase = 1.0961 (in radians)
= 62.802 degrees
= 4.19 hours
= 4 h 11 min

According to p-values, we reject the null hypothesis for MESOR, amplitude and acrophase at 5% level of significance. Since, p-values are less than alpha = 0.05 ie., 5% level of significance. Hence, MESOR, amplitude and acrophaseare statistically significant from zero.



Also, the 95% confidence interval of the amplitude, shown at the peak, does not overlap the MESOR hence the amplitude of the fitted curve is significantly greater than zero.

For Apoorv's Diastolic BP

Data consists of 420 readings, but 17 are the missing values. Hence, we have 403 readings with corresponding times at which readings were taken.

Since it is a circadian rhythm, we have taken the period to be 24 hours.

Following is the summary for the cosinor fit to the readings :

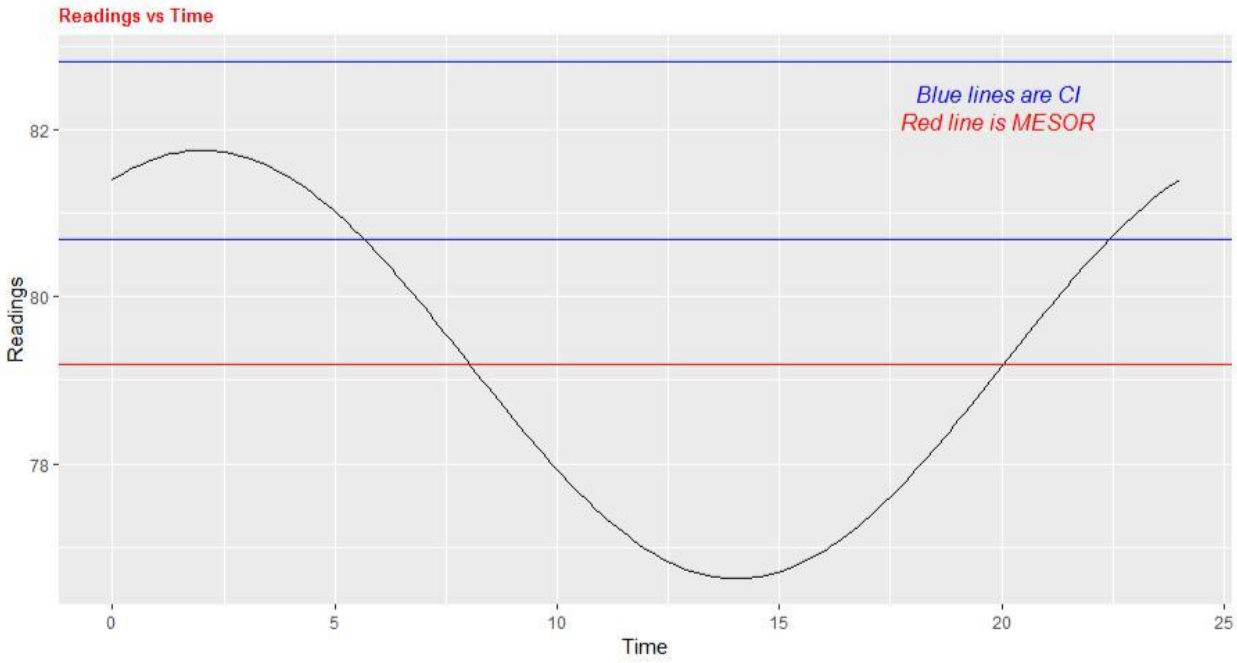
```
Raw model coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept)   79.1896         0.3414  78.5206  79.8587 0.0000
rrr            2.2114         0.5125   1.2069   3.2160 0.0000
sss            1.3063         0.4664   0.3921   2.2206 0.0051

**********************

Transformed coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept)   79.1896         0.3414  78.5206  79.8587 0.0000
amp            2.5684         0.5444   1.5015   3.6354 0.0000
acr            0.5336         0.1670   0.2063   0.8608 0.0014
```

Interpretations:

From the above summary,
MESOR is the intercept = 79.1896

Amplitude = 2.5684
Acrophase = 0.5336 (in radians)
= 30.573 degrees
= 2.04 hours
= 2 h 2 min

According to p-values, we reject the null hypothesis for MESOR, amplitude and acrophase at 5% level of significance. Since, p-values are less than alpha = 0.05 ie., 5% level of significance. Hence, MESOR, amplitude and acrophaseare statistically significant from zero.



Also, the 95% confidence interval of the amplitude, shown at the peak, does not overlap the MESOR hence the amplitude of the fitted curve is significantly greater than zero.

6.3.3 For Akash'sBlood pressure

For Akash's Systolic BP

Data consists of 420 readings, but 11 are the missing values. Hence, we have 409 readings with corresponding times at which readings were taken.

Since it is a circadian rhythm, we have taken the period to be 24 hours.

Following is the summary for the cosinor fit to the readings :

```
Raw model coefficients:
estimatestandard.error lower.CI upper.CI p.value
```

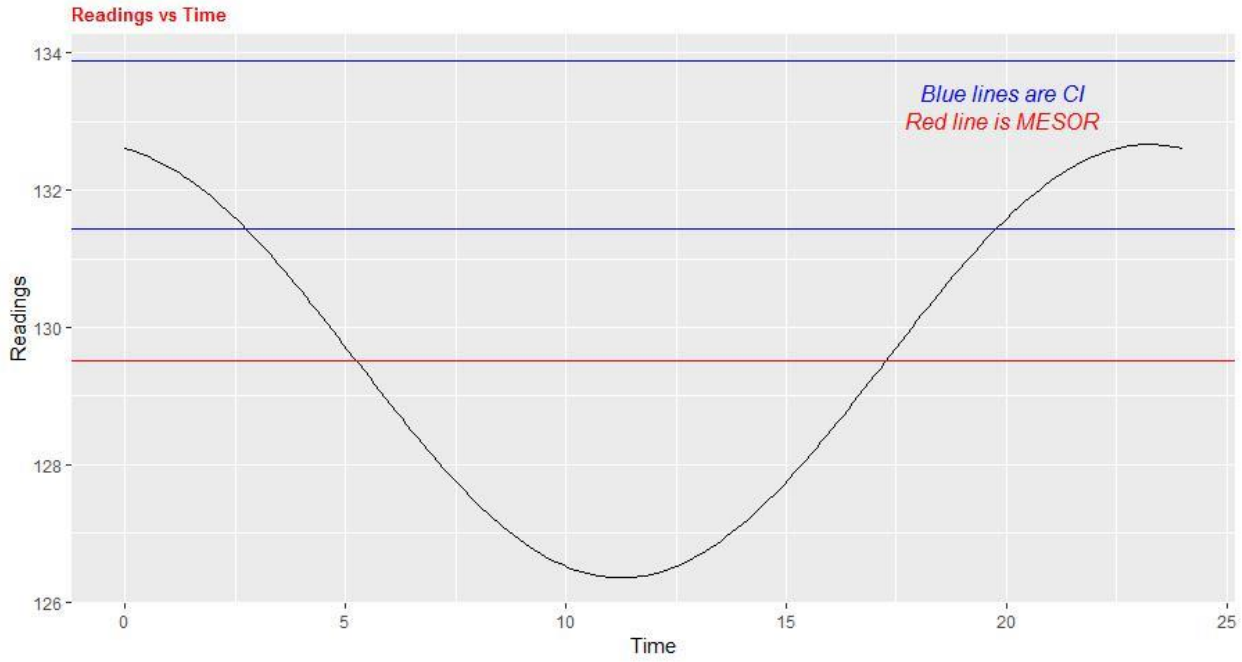
```
(Intercept) 129.5085          0.4352 128.6556  130.3614 0.0000
rrr           3.0960          0.6515   1.8191    4.3729 0.0000
sss          -0.6127          0.5953  -1.7795    0.5540 0.3034

**********************

Transformed coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept) 129.5085          0.4352 128.6556  130.3614  0.000
amp           3.1560          0.6241   1.9328    4.3793  0.000
acr          -0.1954          0.1977  -0.5829    0.1921  0.323
```

Interpretations:

From the above summary,
MESOR is the intercept = 129.5085
Amplitude = 3.1560
Acrophase = -0.1954 (in radians)
= -11.196 degrees
= -0.75 hours = 45 min
= 23 h 15 min

According to p-values, we reject the null hypothesis for MESOR and amplitude at 5% level of significance. Since, p-values are less than alpha = 0.05 ie., 5% level of significance. Hence, MESOR and amplitude are statistically significant from zero.

But, acrophase is insignificant which means peak is at initial time point.



Also, the 95% confidence interval of the amplitude, shown at the peak, does not overlap the MESOR hence the amplitude of the fitted curve is significantly greater than zero.

For Akash'sDiastolic BP

Data consists of 420 readings, but 11 were the missing values. Hence, we have 409 readings with corresponding times at which readings were taken.

Since it is a circadian rhythm, we have taken the period to be 24 hours.

Following is the summary for the cosinor fit to the readings :

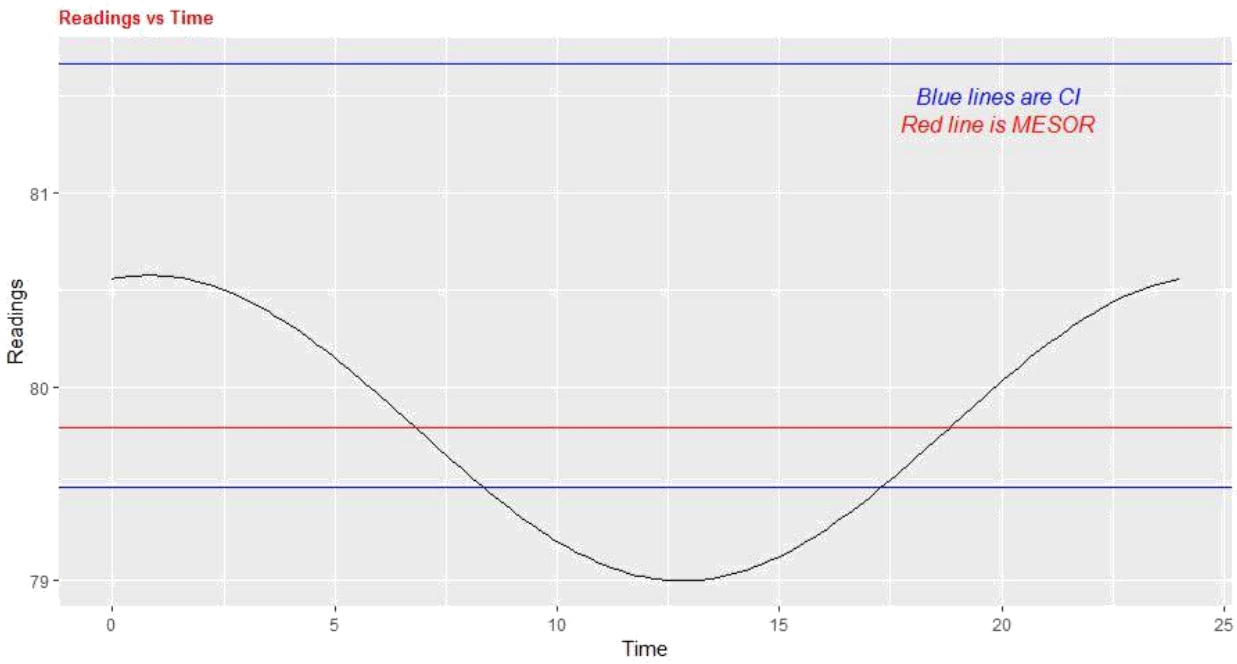
```
Raw model coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept)   79.7873         0.3587  79.0842  80.4904 0.0000
rrr            0.7687         0.5370  -0.2839   1.8213 0.1523
sss            0.1677         0.4907  -0.7941   1.1295 0.7325

**********************

Transformed coefficients:
estimatestandard.error lower.CI upper.CI p.value
(Intercept)   79.7873         0.3587  79.0842  80.4904 0.0000
amp            0.7868         0.5570  -0.3049   1.8784 0.1578
acr            0.2148         0.5948  -0.9510   1.3806 0.7180
```

Interpretations:

From the above summary,
MESOR is the intercept = 79.7873
Amplitude = 0.7868
Acrophase = 0.2148 (in radians)
= 12.307 degrees
= 0.82 hours
= 49 min

According to p-values, we do not reject the null hypothesis for MESOR, amplitude and acrophase at 5% level of significance. Since, p-values are greater than alpha = 0.05 ie., 5% level of significance. Hence, MESOR, amplitude and acrophaseare statistically insignificant.

Amplitude is insignificant means that there is no rhythmicity in the data. As we know, amplitude is length between MESOR and peak and being insignificant it implies there is no significant difference between MESOR and amplitude.



Also, the 95% confidence interval of the amplitude, shown at the peak, overlaps the MESOR which means the amplitude of the fitted cu

7. Conclusions

- All the three patients have significant effect of time on their systolic as well as diastolic blood pressure except diastolic blood pressure of Akash.

- For all patients Systolic blood pressure gives 24 hours of rhythmicity (Not significant only for Apoorv;sSys.blood pressure)

- For all patients Diastolic blood pressure does not give 24 hours rhythmicity.

- MESOR. amplitude and acrophase are significant for Mithilesh's systolic blood pressure, Apoorv's systolic blood pressure and Apoorv's diastolic blood pressure. They infer 24 hour rhythmicity.

- While none are significant for Akash's diastolic blood pressure implies that Akash's circadian rhythm is imbalanced.

- Acrophase are insignificant for Mithilesh's diastolic blood pressure and Akash's systolic blood pressure implies they have higher readings of diastolic blood pressure at initial time point i.e., at start of new day means that at 24th hour each day.

8. References

1)
ROBERTO REFINETTI, GERMAINE CORNÉ LISSEN, and FRANZ HALBERG. (2003).
Procedures for numerical analysis of circadian rhythms. Circadian Rhythm Laboratory,South Carolina, SC, USA and Chronobiology Center, University of Minnesota, MN, USA: Biol Rhythm Res. 2013 ; 38(4): 275–325. doi:10.1080/09291010600903692.

2)
F. A. M. Frescura, C. A. Engelbrecht and B. S. Frank. (2008, August 5). *academic.oup.* Retrieved from https://academic.oup.com/mnras/article/388/4/1693/981666

3)
Gierke, C. L. (2013, July). *ANALYSIS OF RHYTHMS USING R: CHRONOMICS ANALYSIS TOOLKIT (CAT)*. Retrieved from https://564394709114639785.weebly.com/uploads/4/6/0/5/4605921/analysis_of_rhythms_in_r-gierke_2013.pdf

Links

(n.d.).Retrieved from https://www.pathwayz.org/Tree/Plain/BIOLOGICAL+RHYTHMS+-PHASE+SHIFTS.

(n.d.).Retrieved from https://www.circadian.org/dictionary.html.

(n.d.).Retrieved from https://564394709114639785.weebly.com/cosinor.html.

9. Appendix

9.1 For Lomb ScarglePeriodogram

For constructing Lomb ScarglePeriodogram,we have used “lomb” package in “R”

We have used lsp function in R.

lsp(x, times = NULL, from = NULL, to = NULL, type = c("frequency", "period"), ofac = 1, alpha = 0.01, plot = TRUE, ...)

where arguments are as follows :

x	The data to be analysed. x can be either a two-column numerical dataframe or matrix, with sampling times in columnn 1 and measurements in column 2.
times	If x is a single vector, times can be provided as a numerical vector of equal length containing sampling times. If x is a vector and times is NULL, the data are assumed to be equally sampled and times is set to 1:length(x).
from	The starting frequency (or period, depending on type) to begin scanning for periodic components.
to	The highest frequency (or period, depending on type) to scan.
type	Either “frequency” (the default) or “period”. Determines the type of the periodogram x-axis.
ofac	The oversampling factor. Must be an integer>=1. Larger values of ofac lead to finer scanning of frequencies but may be time-consuming for large datasets and/or large frequency ranges (from...to).
alpha	The significance level. The periodogram plot shows a horizontal dashed line. Periodogram peaks exceeding this line can be considered significant at alpha. Defaults to 0.01. Only used if plot=TRUE.
plot	Logical. If plot=TRUE the periodogram is plotted.

Following are the codes:

```
d=read.table(file.choose(),header=TRUE,sep=",")          #Extracting data

attach(d)                                  #Attaching data to extract columns

Sys=d[,2]                                  # second column for systolic bp
Dia=d[,3]                                   #third column for diastolic bp

time1=d$Time                                                  #column of time
```

```
length(Time)                                    #length of time

LSP=lsp(Sys,time1,type="period",from=6,to=1439,ofac=58)
#lsp for systolic bp

LSO=lsp(Dia,time1,type="period",from=6,to=1439,ofac=58)
 #lsp for diastolic bp

summary(LSP)                          #summary of lsp for systolic bp
summary(LSO)                                  #summary of lsp for diastolic bp
```

9.2For CosinorPeriodogram

For constructing cosinorperiodogram,we have used “cosinor” package in “R”

We have used cosinor function in R.

cosinor.lm(formula, period = 12, data, na.action = na.omit)

where arguments are as follows :

formula	formula specifying the model. Indicate the time variable with time()and covariate effects on the amplitude and acrophase with amp.acro
period	Length of time for a complete period of the sine curve.
data	Data frame where variable can be found
na.action	What to do with missing data

Following are the codes:

```
D=read.table(file.choose(),header=TRUE,sep=",") #Extracting dataset
View(D) #Viewing dataset
tail(D) #last six observations of the dataset
D=D[,c(1,2,3)]     #Extracting three columns
dim(D)        #dimension of the dataset
names(D)   #column names

Time=D[,2]   #second column as time
Y=D[,3]      #third column is blood pressure
Y=as.numeric(Y)   #taking as numerical values
Time=as.numeric(Time) #taking as numerical values
```

```
N=dim(D)[1]   #number of rows

p=c()        #vector for predictions
m=list()     #list for mean square errors
library(cosinor)   #calling library
for(i in 1:24)     #loop for each period from 1 to 24
{
fit=cosinor.lm(Y~time(Time), period =i, data=D ,na.action=na.omit)   #fitting cosinor


p=predict(fit) #predictions m[[i]]=mean((Y-
p)^2) #mean square errors }


m=unlist(m)                                                    #unlisting

d=data.frame(1:24,m)     #data frame for mse with corresponding periods
d=as.matrix(d)          #convertinbg to matrix form
d=d[order(d[,2]),]      #ordering(ascending) periods according to mse

d1=as.data.frame(d) #converting to data frame
head(d1) #first six observations

#saving external file for mse
write.table(d1,file="C:\\Users\\admin\\Documents\\Datasets\\Timing in hours without
missing imputed\\MSE\\Akash_Dia.csv",sep=",")


index=c(which(m==min(m))) #period corresponding to minimum mse
which(is.na(Y)==TRUE) #finding missing values

fit=cosinor.lm(Y~time(Time), period =24, data=D
,na.action=na.omit) #cosinor fit with period 24 hour

summary(fit)       #summary of the fit

mesor=fit$coefficients[1]                    #MESOR
amp=fit$coefficients[2]                      #Amplitude
```

```
acro=fit$coefficients[3]                    #Acrophase
```

```
p1=predict(fit)     #predictions

library(ggplot2)                                    #calling library
g=ggplot.cosinor.lm(fit, x_str = NULL)+labs(title="Readings vs Time", y="Readings",
x="Time")
g#plot of cosinor fit

#plotting 95% confidence interval of amplitude at peak
g=g+theme(plot.title=element_text(size=10,face="bold",color="red"))+
geom_hline(yintercept = mesor-0.3049 ,col="blue")+
geom_hline(yintercept = mesor + 1.8784 ,col="blue")+
geom_hline(yintercept = mesor,col="red")

library(grid))                                      #calling library
my_text= "Blue lines are CI"
my_text2= "Red line is MESOR"
my_grob = grid.text(my_text, x=0.8,  y=0.9, gp=gpar(col="Blue", fontsize=12,
fontface="italic"))
my_grob2 = grid.text(my_text2, x=0.8, y=0.85, gp=gpar(col="red", fontsize=12,
fontface="italic"))
h4=g + annotation_custom(my_grob) +annotation_custom(my_grob2)
print(h4)                                           #plot of cosinor fit

a=aggregate(p1~Time,FUN=mean)
 #Splits the data into subsets with respect to time and taking mean over each time

plot(a,type='b',ylab="Diastolic BP",col="red")
#plotting those means(predicted values)

b=aggregate(Y~Time,FUN=mean)
#Splits the data into subsets with respect to time and taking mean over each time

lines(b,col="blue")      #plotting means on same graph (actual values)

legend(locator(1),legend=c("Actual","cosinor/predicted"),lty=1:2,col=c("blue","red"),cex=
0.8)


plot(Y,p1,xlab="Actual",ylab="cosinor/predicted")
```

```
#plotting actual vs predicted values

summary(fit)  #summary of fit with period of 24 hrs


amp=fit$coefficients[2]     #Amplitude
acr=fit$coefficients[3]       #Acrophase
M=fit$coefficients[1]         #Amplitude

#phi(or acr)=x*(2*pi)/lambda
#lambda=total wavelength = period

acro=(24*(acr))/(2*pi)   #Acrophase = (acr * period) / (2 * 3.14)
acro#length (acrophase)

#95%  CI

beta=c()
beta=(2)*mean(cos((2*pi*Time)/24)*Y)                          #beta coefficient

gamma=c()
gamma=(2)*mean(sin((2*pi*Time)/24)*Y)                         #gamma coefficient


acr=-180-atan(abs(gamma/beta))                                #Acrophase

ssq1=c()
ssq1=sum(Y-(M+amp*(cos(acr)*cos(2*pi*Time/24)- sin(acr)*sin(2*pi*Time/24))))^2/(N-
3)                              #Residual variance

(N*(beta^2+gamma^2))/(4*ssq1)                                 #F calculated value

qf(0.95,2,N-3)                                                #F tabulated value
```