John R. Zaleski, Ph.D., CPHIMS
john@medicinfotech.com

Investigating the use of the Lomb-Scargle Periodogram for Heart Rate Variability Quantification

The use of the Lomb-Scargle Periodogram (LSP) for the analysis of biological signal rhythms has been well-documented. [1,2]

> *"The analysis of time-series of biological data often require special statistical procedures to test for the presence or absence of rhythmic components in noisy data, and to determine the period length of rhythms."*[3]

> *"In the natural sciences, it is common to have incomplete or unevenly sampled time series for a given variable. Determining cycles in such series is not directly possible with methods such as Fast Fourier Transform (FFT) and may require some degree of interpolation to fill in gaps. An alternative is the Lomb-Scargle method (or least-squares spectral analysis, LSSA), which estimates a frequency spectrum based on a least squares fit of sinusoid."*[4]

> *"The standard way to deal with such data is to compute a discrete Fourier transform of the data and then view the transform as an absorption spectrum, a power spectrum, a Schuster periodogram (Schuster 1905), or a Lomb-Scargle periodogram (Lomb 1976, and Scargle 1982 and 1989), see Priestley (1981) and Marple 1987 for a review of classical spectral estimation techniques. The problem with all such techniques is that they have not be derived from any single set of unifying principles that tell one what is the optimal way to estimate the period. In this paper we change that by using Bayesian probability theory to deriving the discrete Fourier transform, the power spectrum, the weighted power spectrum, the Schuster periodogram and the Lomb-Scargle periodogram as special cases of a generalized LombScargle periodogram, and show that the generalized Lomb-Scargle periodogram is a sufficient statistic for single frequency*

---

[1] T. Ruf, "The Lomb-Scargle Periodogram in Biological Rhythm Research: Analysis of Incomplete and Unequally Spaced Time-Series." *Biological Rhythm Research*, 1999, Vol. 30, No. 2, pp. 178-201.
[2] Jozef Púčik, "Heart Rate Variability Spectrum: Physiologic Aliasing and Nonstationarity Considerations." *Trends in Biomedical Engineering*. Bratislava, September 16-18, 2009.
[3] T. Ruf, "The Lomb-Scargle Periodogram in Biological Rhythm Research: Analysis of Incomplete and Unequally Spaced Time-Series". Biological Rhythm Research, 1999, Vol. 30, No. 2, pp. 178-201.
[4] *Marc in the box,* "Lomb-Scargle periodogram for unevenly sampled time series"
", http://www.r-bloggers.com/lomb-scargle-periodogram-for-unevenly-sampled-time-series/.
Published January 10th, 2013. Accessed 20-April-2015.

> *estimation, (a suffi- cient statistic summarizes all of the information in the data relevant to the question being asked).”*[5]

A key benefit of the LSP is in its ability to process unequally-spaced or missing data in terms of time. This is a distinct advantage over the discrete time Fourier Transform (DTFT). The normalized power is determined from:

$$PN(\omega) = \tfrac{1}{2\sigma^2}\left\{ \frac{\left[\sum_j \left(Y_j - \bar{Y}\right)\cos\omega(t_j - \tau)\right]^2}{\sum_j \cos^2\omega(t_j - \tau)} + \frac{\left[\sum_j \left(Y_j - \bar{Y}\right)\sin\omega(t_j - \tau)\right]^2}{\sum_j \sin^2\omega(t_j - \tau)} \right\}$$, and

$\tau = \left(\frac{1}{2\omega}\right)\tan^{-1}\left[\frac{\sum_j \sin 2\omega t_j}{\sum_j \cos 2\omega t_j}\right]$, where, $PN(\omega)$ is the normalized power as a function of angular frequency. The relationship between angular (or circular) frequency and frequency in Hz is given by:

$$f = \frac{\omega}{2\pi}.$$

The solution of these equations is rather straightforward. Let’s take a look at a simple example and then extrapolate to a more complex case.

Consider following sine wave given in the following diagram. This sine wave is described by the equation:

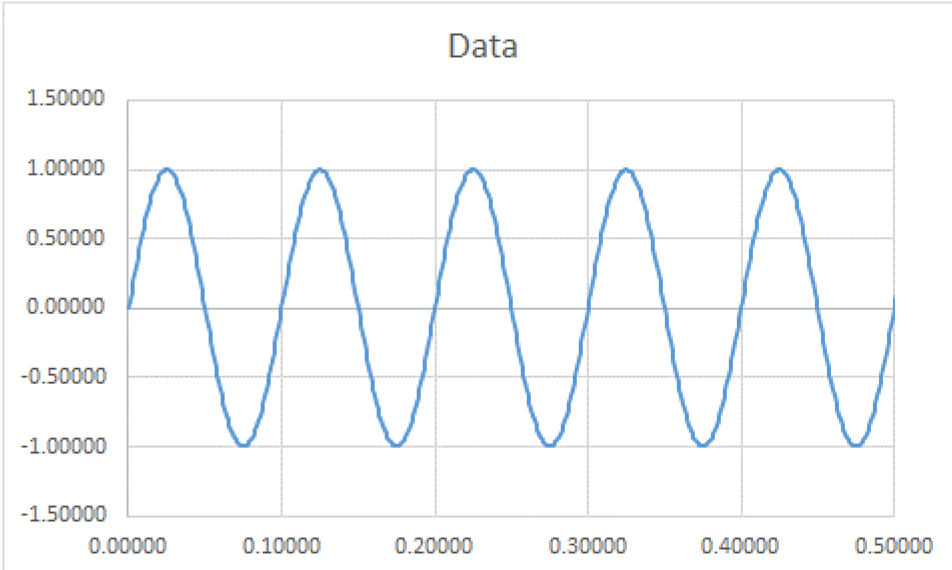
$y(t) = A\sin\omega t$, where A is amplitude (assumed to be 1), ω is the radian or circular frequency, and t is time.

The relationship between circular frequency and frequency in Hertz is given above.

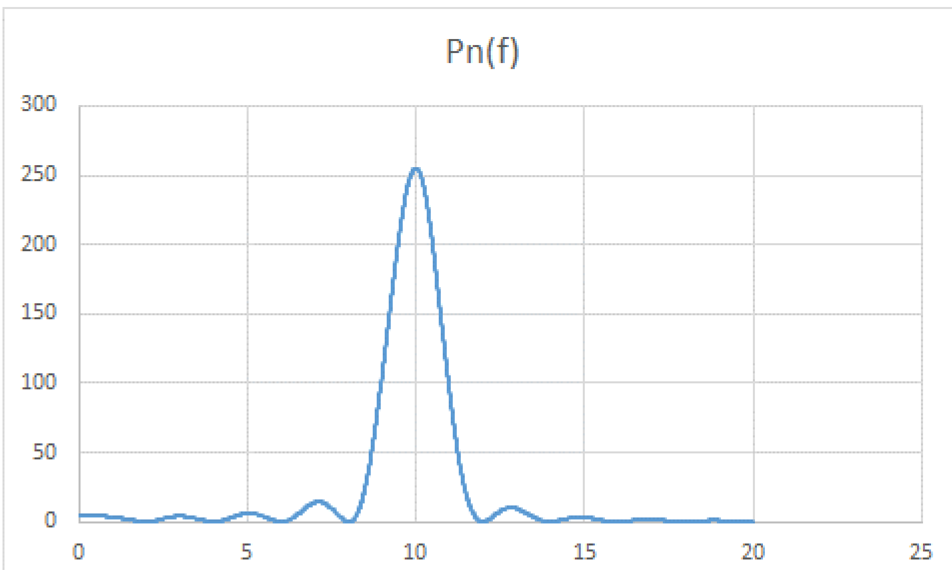
In the plot of the following figure $f = 10Hz$. Hence, the sine equation becomes:

$$y(t) = \sin(2\pi \times 10t) = \sin(20\pi t).$$

[5] *G. Larry Brethorst*, “Frequency Estimation And Generalized Lomb-Scargle Periodograms” http://bayes.wustl.edu/glb/Lomb1.pdf. Accessed 20-April-2015.

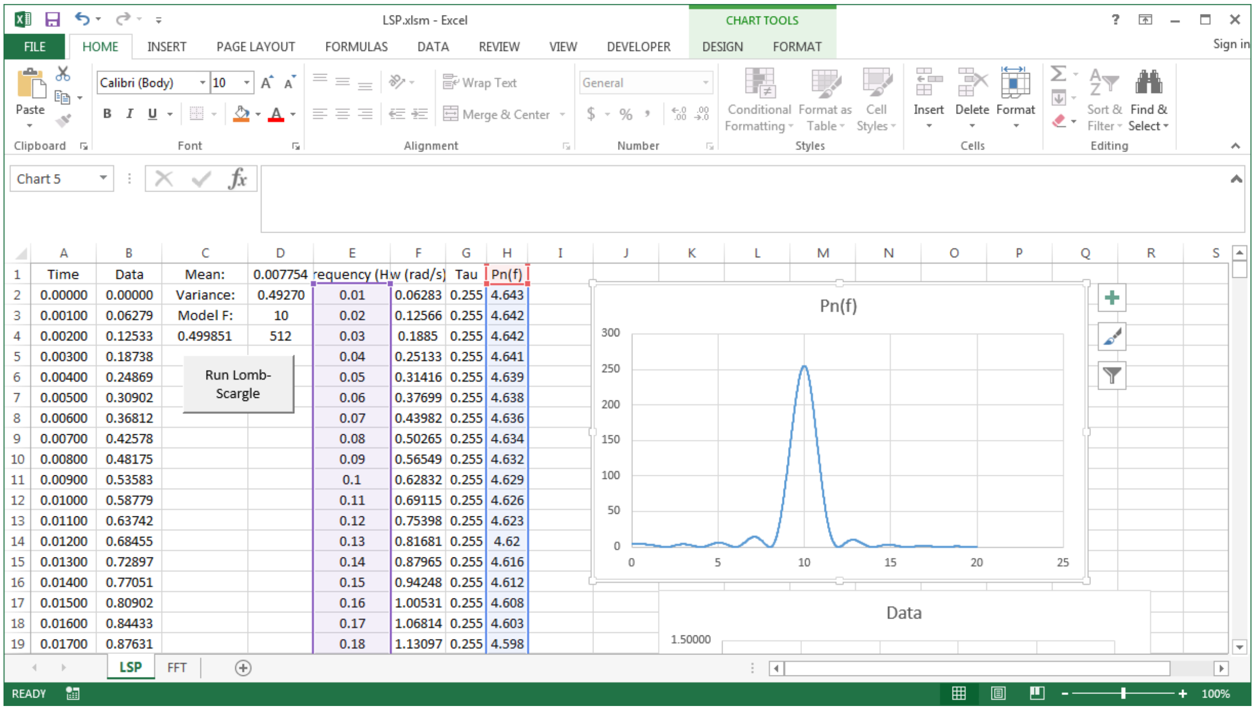


The Lomb-Scargle Periodogram for this simple waveform is given by the following plot:



The maximum is centered about the principle frequency of 10Hz. The method used to calculate the LSP is written in Visual Basic and incorporated as a Macro into an Excel spreadsheet. The spreadsheet is shown in the following figure.

The code used to create the LSP plot above is included in the table on the next page.



Visual Basic code listing is as follows:

```
Sub CalculateLS()
'
' Program to calculate the Lomb-Scargle Periodogram
' of a time-based signal in 1 dimension.
'
' Copyright(c) 2015
' John R. Zaleski, Ph.D., CPHIMS
'
' Free to use with proper attribution to the author.
'
Dim Pi As Double
Dim s As Double
Dim c As Double
Dim mean As Double
Dim numRows As Integer
Dim rowCounter As Integer
Dim omegaRow As Integer
Dim Ybar As Double
Dim variance As Double
Dim n1 As Double
Dim n2 As Double
Dim d1 As Double
Dim d2 As Double
Dim tj As Double
Dim omega As Double
Dim f As Double
Dim tau As Double
Dim Maxf As Double
Dim Deltaf As Double
Dim fRow As Integer

'
' Clear contents of output columns and set headers
'
```

```
Worksheets("LSP").Columns(5).EntireColumn.Clear
Worksheets("LSP").Columns(6).EntireColumn.Clear
Worksheets("LSP").Columns(7).EntireColumn.Clear
Worksheets("LSP").Columns(8).EntireColumn.Clear

Worksheets("LSP").Cells(1, 5) = "Frequency (Hz)"
Worksheets("LSP").Cells(1, 6) = "w (rad/s)"
Worksheets("LSP").Cells(1, 7) = "Tau"
Worksheets("LSP").Cells(1, 8) = "Pn(w)"

'
' Declare value for Pi
'

Pi = 3.1415926535

'
' Determine number of rows of data in columns
' 1 & 2
'

rowCounter = 1
While Worksheets("LSP").Cells(rowCounter, 1) <> ""
    rowCounter = rowCounter + 1
Wend

'
' Define the number of processing rows
' as excluding the header, and assuming
' the first value begins at count of zero,
' not one.
'

numRows = rowCounter - 2

'
' Calculate mean of signal,
' column 2. Assign value to
' column 4, row 1.
'

omegaRow = 2
mean = 0
For Row = 2 To rowCounter
    mean = mean + Worksheets("LSP").Cells(Row, 2)
Next
mean = mean / numRows
Worksheets("LSP").Cells(1, 4) = mean

'
' Calculate variance of signal,
' column 2. Assign value to
' column 4, row 2.
'

variance = 0
For Row = 2 To rowCounter
    variance = variance + (Worksheets("LSP").Cells(Row, 2) - mean) ^ 2
Next
variance = variance / (numRows - 1)
Worksheets("LSP").Cells(2, 4) = variance

'
' Model frequency is contained in
' column 4, row 3.
'
' Assign number of data points to
' columen 4, row 4.
'
```

```
Worksheets("LSP").Cells(4, 4) = numRows

'
' Create the frequency column, defined by
' the sample frequency of the model in
' cell (3,4).
' Make the maximum frequency twice the
' frequency in this cell to provide
' a better visual spectrum.
' Make the increment in frequency
' something large so as to
' provide better resolution
' for graphing.
'
Maxf = Worksheets("LSP").Cells(3, 4)
Deltaf = Maxf / 1000#
Maxf = Maxf * 2#

'
' Create the frequency column.
' Initialize the first frequency as
' the frequency increment--cannot
' make zero, else singularity.
'
f = Deltaf
fRow = 2
While f < Maxf
    Worksheets("LSP").Cells(fRow, 5) = f
    f = f + Deltaf
    fRow = fRow + 1
Wend

'
' Calculate
'

While Worksheets("LSP").Cells(omegaRow, 5) <> ""

    '
    ' The frequency in Hz is the independent variable.
    ' This could be set up as a for loop instead.
    '

    f = Worksheets("LSP").Cells(omegaRow, 5)

    '
    ' Compute omega. Assign to column 6.
    '

    omega = 2 * Pi * f
    Worksheets("LSP").Cells(omegaRow, 6) = omega
    s = 0
    c = 0

    '
    ' Compute denominators
    '

    For Row = 2 To numRows
        tj = Worksheets("LSP").Cells(Row, 1)
        s = s + Sin(2 * omega * tj)
        c = c + Cos(2 * omega * tj)
    Next

    '
    ' Compute tau. Assign to column 7.
    '

    tau = 1 / (2 * omega) * Atn(s / c)
```

```
    Worksheets("LSP").Cells(omegaRow, 7) = tau

    '
    ' Initialize PN(w) summations.
    '

    n1 = 0
    n2 = 0
    d1 = 0
    d2 = 0

    '
    ' Compute PN(w). Assign to column 8.
    '

    For Row = 2 To numRows
        tj = Worksheets("LSP").Cells(Row, 1)
        yj = Worksheets("LSP").Cells(Row, 2)
        n1 = n1 + (yj - Ybar) * Cos(omega * (tj - tau))
        n2 = n2 + (yj - Ybar) * Sin(omega * (tj - tau))
        d1 = d1 + (Cos(omega * (tj - tau))) ^ 2
        d2 = d2 + (Sin(omega * (tj - tau))) ^ 2
    Next

    PN = (1 / (2 * variance)) * ((n1 * n1) / d1 + (n2 * n2) / d2)
    Worksheets("LSP").Cells(omegaRow, 8) = PN

    '
    ' Increment row. Grab next frequency.
    '
    omegaRow = omegaRow + 1
Wend

'
' Center align all output columns
'

Worksheets("LSP").Columns("E").ColumnWidth = 10
Worksheets("LSP").Columns("F").ColumnWidth = 7
Worksheets("LSP").Columns("G").ColumnWidth = 5
Worksheets("LSP").Columns("H").ColumnWidth = 5

Worksheets("LSP").Columns("E").VerticalAlignment = xlVAlignCenter
Worksheets("LSP").Columns("E").HorizontalAlignment = xlHAlignCenter
Worksheets("LSP").Columns("F").VerticalAlignment = xlVAlignCenter
Worksheets("LSP").Columns("F").HorizontalAlignment = xlHAlignCenter
Worksheets("LSP").Columns("G").VerticalAlignment = xlVAlignCenter
Worksheets("LSP").Columns("G").HorizontalAlignment = xlHAlignCenter
Worksheets("LSP").Columns("H").VerticalAlignment = xlVAlignCenter
Worksheets("LSP").Columns("H").HorizontalAlignment = xlHAlignCenter

End Sub
```

The program may be used to study behavior of time-varying signals, such as heart rate (i.e., R-R interval) or other signals which may possess uneven time intervals or missing data.

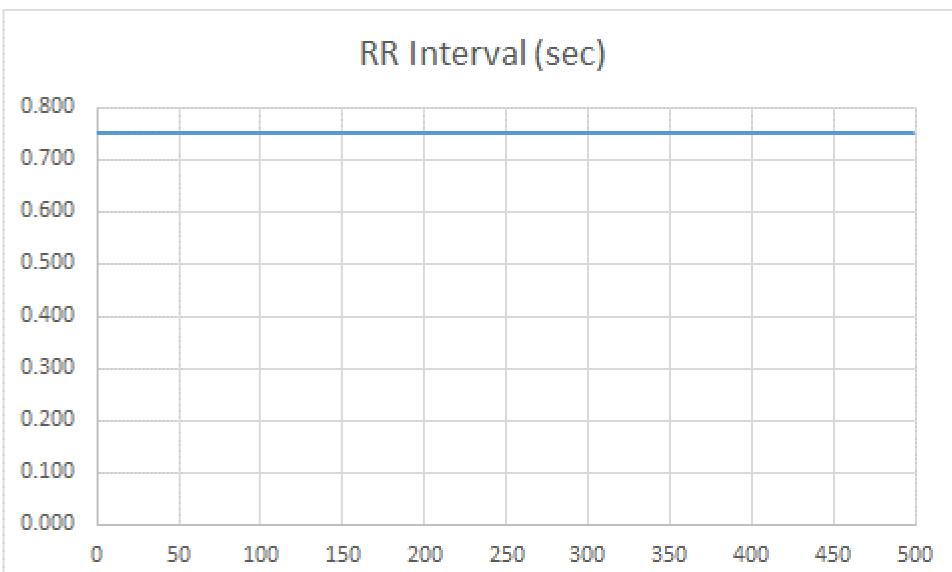
Heart rate is calculated from two identical points on the standard electrocardiogram (ECG or EKG). The calculation is such:

$HR = \frac{60}{\Delta t_{RR}}$, where Δt_{RR} is the R-R interval. For a heart rate of 80 beats per minute, the R-R interval is computed to be: $\Delta t_{RR} = \frac{60}{80} = 0.750s = 750m\sec$

While normal sinuous rhythm is usually between 60 and 100 beats/minute, in patients with problems, the R-R interval can vary. For example, an acceleration of heart rate with inspiration and a slowing with expiration can be representative of an arrhythmia.

Heart Rate Variability (HRV) has been used as an assessment of the autonomic nervous system, based on sympathetic and parasympathetic tone (SNS versus PSNS).[6] High HRV is indicative of parasympathetic tone. Low HRV is indicative of sympathetic tone. Low HRV has been associated with coronary heart disease and those who have had heart attacks. A low HRV among those who have had heart attacks places those individuals at higher risk of death within 3 years.[7]

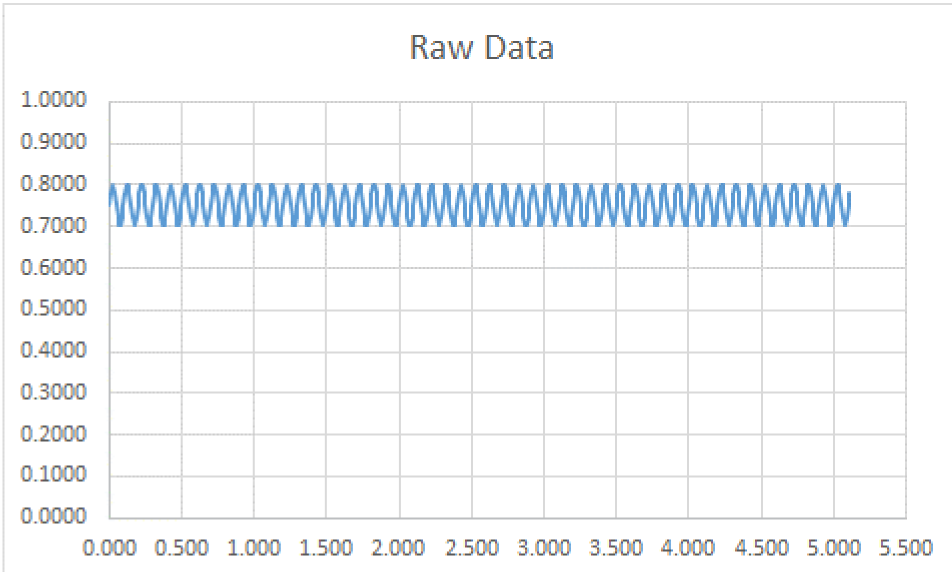
We can look at some simulated data to gain a better understanding of the measurement of HRV and its visualization. The following plot shows a constant heart rate of 80 beats/minute, as represented by the constant RR interval of 750 milliseconds. The plot shows the RR interval (vertical axis) as a function of measurement number (horizontal axis).



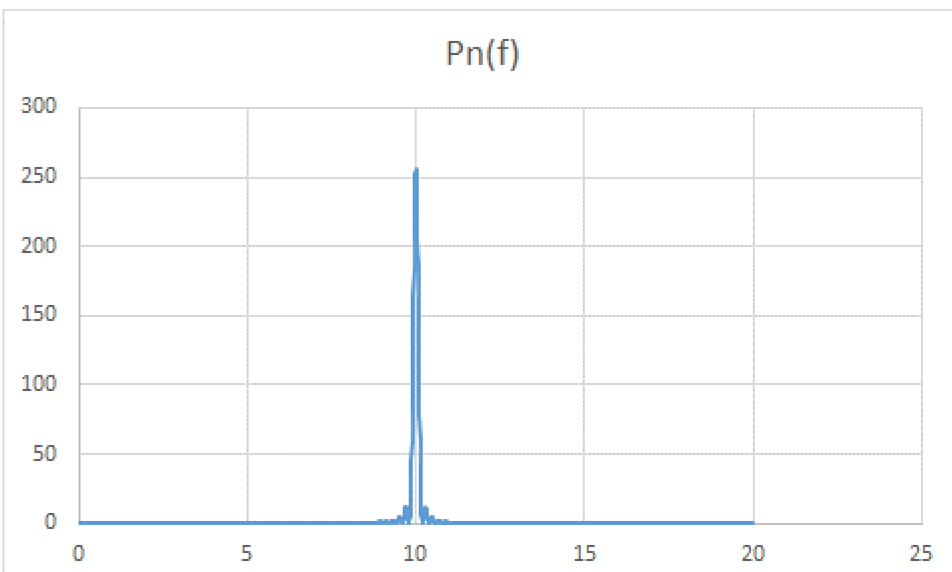
For a constant signal, as per above, the LSP is non-existent (i.e., no periodic behavior).

[6] http://www.megaemg.com/knowledge/heart-rate-variability-hrv/ Accessed 21-April-2015
[7] "Mark's Daily Apple", http://www.marksdailyapple.com/have-you-checked-your-heart-rate-variability-lately/#axzz3XxaoLwNt Accessed 21-April-2015

When an artificial amount of variability, corresponding to sinusoidal perturbative behavior, having frequency equal to 10 Hz, is added onto the signal, as shown in the following figure, the oscillatory behavior is overlaid on the constant offset:

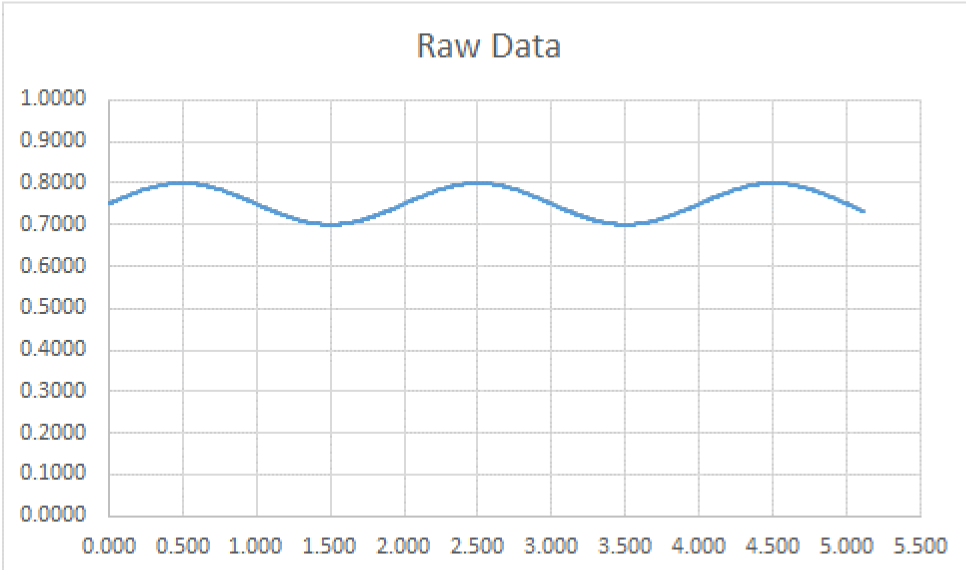


When the mean signal is subtracted from the oscillatory component, the following LSP is computed:

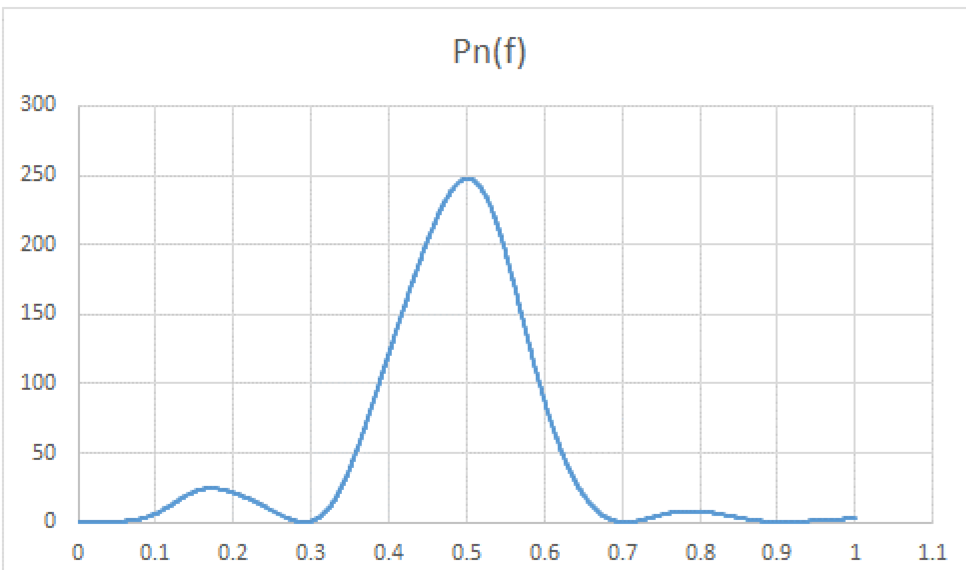


The plot in this figure illustrates that the oscillatory component is isolated and shown to be 10 Hz, verifying the actual oscillatory component of the perturbative signal.

As the frequency is reduced, this is visible in the raw data plot, wherein f = 0.5 Hz, per the plot below:



The corresponding LSP become:

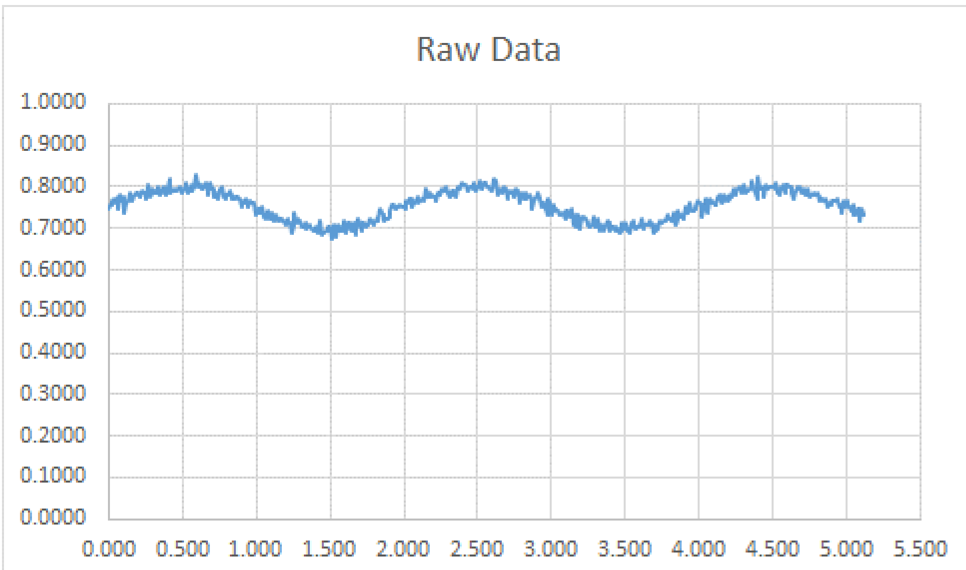


Note, again, that the constant (or D.C.) component has been removed by subtracting the mean value of the raw signal. Hence,

$y(t) = y_R(t) - \bar{y}$, where $\bar{y}$ is the mean or average value of the signal. If we leave the constant component in, we will see a large value of the power expressed at a frequency of zero. If the amplitude of the perturbation is far less in magnitude than the constant component, then the power of the constant component will be very large when depicted in the power plot as above and the frequency component will not be as visible. For example, suppose a model of the heart rate is represented as follows:

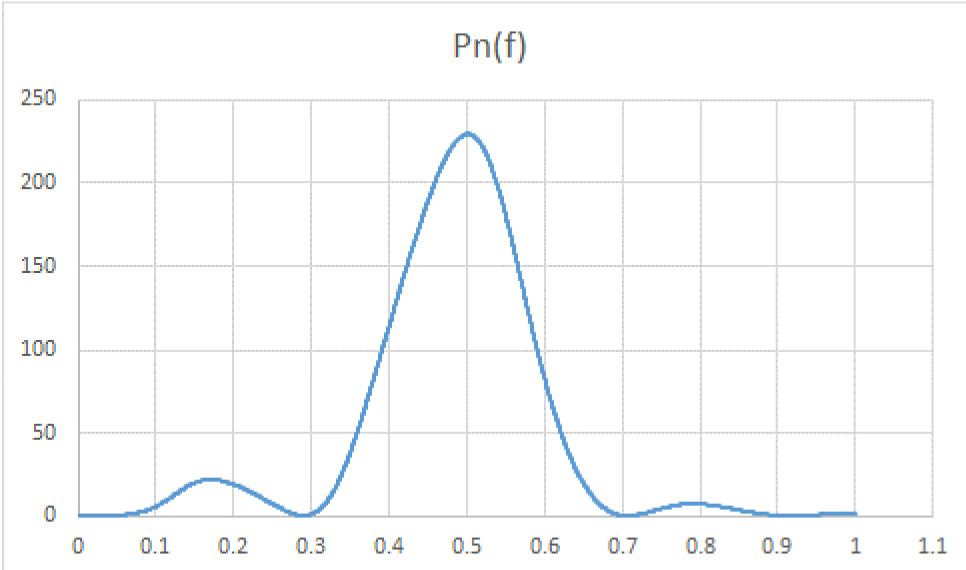
$y(t) = HR_{Mean} + HRV_{Pert}\sin(2\pi ft)$, where HR_{Mean} is the average heart rate and HR_{Pert} is the amplitude of the variable or perturbative component. If $HR_{Mean} >> HR_{Pert}$ then the LSP will be dominated at zero frequency and the frequency of the perturbative component will be less easy to see in the plot. Hence, the mean component is subtracted off to reveal only the perturbation.

We can also add a small amount of white noise to the signal, as shown in the following plot. This will have the effect of simulating artifact or randomness:

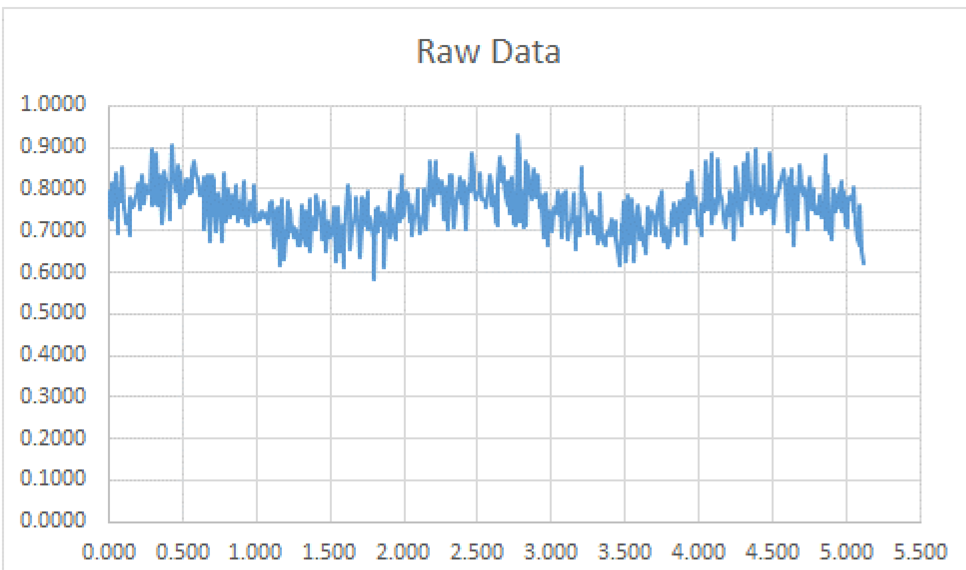


Mathematically, the addition of the white noise to the signal is given by:

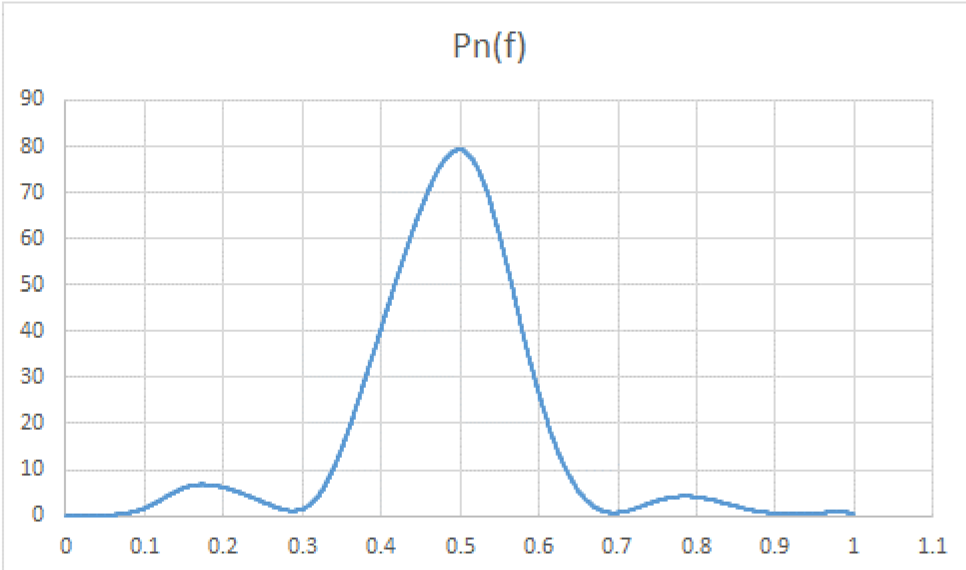
$y(t) = HR_{Mean} + HRV_{Pert}\sin(2\pi ft) + v_{WN}$, where v_{WN} is a random variate drawn from a Gaussian distribution having zero mean and some variance. In the case of the plot above, the variance is 0.01. The white noise is additive on top of the original signal. A corresponding plot of the LSP is as follows, illustrating rather well that the white noise has little effect on the determination of the underlying perturbation frequency:



Even when the noise is increased to 0.05, as shown in the following plot, with a concomitant increase in the noise of the signal, the underlying LSP is still unaffected:



The corresponding LSP is shown in the following plot:



Summary

The preceding analysis demonstrates a viable method for researching the effects of signal variability. The next step is to determine thresholds in variability that correspond to physiological events, and to identify these events relative to the changes in variability determination.