

Optimal RR-Interval Data Length for Entropy based Heart Rate Variability Analysis

Manjit Singh

Dept. of Electronics and
Communication
Guru Nanak Dev University,
Regional Campus,
Jalandhar-144009

Butta Singh

Dept. of Electronics and
Communication
Guru Nanak Dev University,
Regional Campus,
Jalandhar 144009

Gurpreet Singh

Dept. of Electronics and
Communication
Guru Nanak Dev University,
Regional Campus,
Jalandhar-144009

ABSTRACT

Heart Rate Variability (HRV) is defined as the variations between consecutive instantaneous heart rates that occur in the heart as a consequence of a complex internal dynamic balance. Nonlinear analysis of HRV is helpful to assess the cardiac health noninvasively. Approximate Entropy and Sample Entropy are mathematical algorithms to measure the predictability or repeatability with in a time series. This paper compares the approximate entropy and sample entropy on different data lengths, which are 20 minutes, 10 minutes, 5 minutes, 3 minutes and 2 minutes respectively. In addition it has been observed that the measuring time of sample entropy can be reducing beyond 5 minutes.

General Terms

Heart Rate Variability, Approximate Entropy, Sample Entropy.

Keywords

Heart Rate Variability, ECG, Approximate Entropy, Sample Entropy, Data Lengths.

1. INTRODUCTION

Heart rate variability (HRV) is fluctuation of RR intervals around their mean time. The measurement of HRV provides a non-invasive measurement of the sympathetic and parasympathetic activities of Autonomic nervous system (ANS) modulated by the sino-atrial node (SA). Sympathetic activities increase the heart rate and parasympathetic activities decrease the heart rate [1-2].

Traditionally, HRV has been evaluated by linear time domain measure such as Standard deviation of all RR intervals (SDNN), Root mean square of the difference of successive RR interval (RMSSD), 50 counts divided by total number of all RR intervals(pNN50) [3]. Also, by the frequency domain in which power spectral density (PSD) of the RR series is calculated terms of very low frequency (VLF, around 0.0033-0.04HZ), low frequency (LF around 0.04-0.15HZ) and high frequency (HF, around 0.15-0.4HZ) [4].

Afterwards, it has been supposed that analysis of HRV based on the methods of non-linear dynamics might elicit valuable information like assessment of the risk of sudden death. The parameters which have been utilized to measure non-linear properties include, approximate Entropy (ApEn) [5], Sample Entropy (SampEn) [6] and Poincare plot (SD1, SD2) [7]. Where, ApEn and SampEn are mathematical algorithms created to measure the predictability or repeatability with in a time series.

G. H. Lin et al analyzed the results of four measurements of HRV in three minute and five minute and concluded that the measurement of three minute data is not equal to the five minute data. The comparison has been created, which is only based on Time-domain and frequency domain parameters. However, it lacks the results of non-linear domain [8].

The purpose of this paper is to analyze the effect of RR interval data length on entropy based HRV and further, to find out the optimal RR interval data length.

2. MATERIAL AND METHODS

2.1 Material

In this paper the MIT-BIH Arrhythmia Database is used for analysis [9]. The database includes annotations files of 48 Half-Hour ECG recording obtained from 47 subjects but for this particular paper 26 data sets has been utilized. These recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10mV range. Beat annotations were obtained by reference beat, rhythm, and signal quality annotations. Annotations are labels that point to specific locations within a recording and describe events at those locations [10].

2.2 Methods

2.2.1. Approximate Entropy

In the field of HRV, Approximate Entropy has been widely utilized from many decades. It provides information about the complexity of data which is of shorten length. Basically, ApEn is employed for the measurement of irregularity of the signal. Higher the values of ApEn more will be the irregularities and vice versa. For the calculation of ApEn must have prior information about these parameters, which are defined as follows a) The length m , b) The tolerance value r , c) The data length N . ApEn was introduced by Pincus [5]. In which he suggested that the tolerance value r should be of 0.1-0.25 times of the Standard Deviation (SD) of the data, and m must be 1 or 2 for N data lengths. These N data lengths must be vary from 100-5000 data points. For a given signal $u(1), u(2), u(3)... u(N)$, where N is the total number of data points. For computation of ApEn m must be a positive integer where r must be a positive real number. In our case, the value of r is equivalent to 20% of SD and $m=2$. The algorithm of ApEn can be summarized as follows:

- i Form m vectors $X(1)$ to $X(N-m+1)$ defined by

$$X(i) = [u(i), u(i+1), \dots, u(i+m+1)], \quad 1 \leq i \leq N-m+1$$
- ii Define the distance $d[X(i), X(j)]$ between the vectors $X(i)$ and $X(j)$ as the maximum absolute difference their respective scalar components:

$$d[X(i), X(j)] = \max_{1,2,\dots,m} (u(i+k) - u(j+k))$$
- iii Define $v^m(i) =$ no. of such that $d[X(i), X(j)] \leq r$ from which $c_i^m(r)$ can be defined as:

$$c_i^m(r) = \frac{v^m(i)}{N-m+1}$$
- iv Take the natural logarithm of each $c_i^m(r)$ and average it over i

$$\phi^m(r) = \frac{1}{N-m+1} \sum_1^{N-m+1} \ln(c_i^m(r))$$
- v Increase the m to $m+1$ and repeat the steps 1 to 4 and find

$$\phi^{m+1}(r) = \frac{1}{N-m} \sum_1^{N-m} \ln(c_i^{m+1}(r))$$
- vi Finally calculate ApEn for the data of length of N

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r)$$

2.2.2. Sample Entropy

In ApEn each sequence is self-matched for each template. ApEn referred as a biased estimator which depicts more similarity in comparison to the truly present for finite N [11]. Biased refers to the case where expected value is differs from the parameters which are estimated by it. A new term Sample entropy is introduced that reduce the bias which arrives due to miss match of the templates [12]. In this series data is sampled from a continuous process. SampEn is differ from ApEn in the sense that a) In this self matches are not counted b) Template wise approach is not utilise to estimate the condition probabilities [6]. Hence, it requires only one template to determine the match of $m+1$ length. After that it also computes the probability (logarithmic) which is associated with the time series. The algorithm of SampEn can be summarised as follows:

- i Form m vectors $X(1)$ to $X(N-m+1)$ defined by

$$X(i) = [u(i), u(i+1), \dots, u(i+m+1)], \quad 1 \leq i \leq N-m+1$$
- ii Define the distance $d[X(i), X(j)]$ between the vectors $X(i)$ and $X(j)$ as the maximum absolute difference their respectively scalar components

$$d[X(i), X(j)] = \max_{1,2,\dots,m} (u(i+k) - u(j+k))$$
- iii Define $v^m(i) =$ no of such that $d[X(i), X(j)] \leq r$ and $i \neq j$ from which $B_i^m(r)$ can be defined as:

$$B_i^m(r) = \frac{1}{N-m+1} v^m(i) \quad \text{Here } i = 1, 2, \dots, N-m$$

- iv For each $i=1, 2, \dots, N-m$ define

$$B_i^m(r) = \frac{1}{N-m+1} v^m(i)$$

Where $v^{m+1}(i) =$ no of $X(j)$ and $d_{m+1}[X(i), X(j)] \leq r \quad i \neq j$
- v Now define $B^m(r) = \frac{1}{N-m} \sum_1^{N-m} B_i^m(r)$ and

$$A^m(r) = \frac{1}{N-m} \sum_1^{N-m} A_i^m(r)$$
- vi Finally SampEn for a finite length of N can be taken as

$$SampEn(m, r, N) = -\ln \left(\frac{A^m(r)}{B^m(r)} \right)$$

3. RESULTS

Kubious software has been used for the analysis of data. It is an advanced tool for studying the variability of heart beat intervals [13]. ApEn and SampEn have been calculated in this software on different data lengths and the mean and standard deviation of ApEn and SampEn has been pointed in Table 1 and Table 2 respectively.

Approximate Entropy-The means and standard deviation of ApEn measure for different data lengths were shown in Table.1 .The results pointed out that ApEn is highly dependent on record length and is lower than expected for short lengths, as lower the value higher will be the regularity. Further, it lacks relative consistency which results in high relative error among different data lengths.

Table 1. Mean and Standard deviation of Approximate Entropy with Relative Error

S. No.	Length	Approximate Entropy	Relative Error
1	20 minute	1.5274±0.1036	0
2	10 minute	1.4324±0.1065	6.2198
3	5 minute	1.1667±0.1185	23.6153
4	3 minute	0.9625±0.2131	37.1743
5	2 minute	0.7411±0.1928	51.4797

Sample Entropy- The analyzing results of sample entropy were shown in Table 2. The results show that SampEn is largely independent of record length and value of sample entropy increases with lower measuring time.it is only because SampEn does not count self-matches. In addition SampEn shows less relative error among different data lengths which shows more self-similarity in the same time series.

Table 2. Mean and Standard deviation of Approximate Entropy with Relative Error

S. No.	Length	Sample Entropy	Relative Error
1	20 minute	1.7721±0.2121	0
2	10 minute	1.8084±0.2327	2.0485
3	5 minute	1.8769±0.2687	5.9139
4	3 minute	1.9213±0.3481	8.4194
5	2 minute	2.0082±0.4347	13.3232

Comparison of approximate and sample entropy is pointed in Fig 1

Relative error- For approximate and sample entropy relative error of individual data length is calculated w.r.t 20 minute data length. Hence, results obtained are shown in Table 1 and Table 2. Further, in fig 2 comparison of relative error of ApEn and SampEn w.r.t data lengths has been depicted. Hence, relative error of SampEn is less in comparison to the ApEn (as shown in Table 1 and Table 2). So, it is possible to reduce the measurement time of sample entropy up to 2 minute with just 13% of relative error but in case of approximate entropy relative error increases drastically after 10 minute data length and causes up to 51% in case of 2 minute.

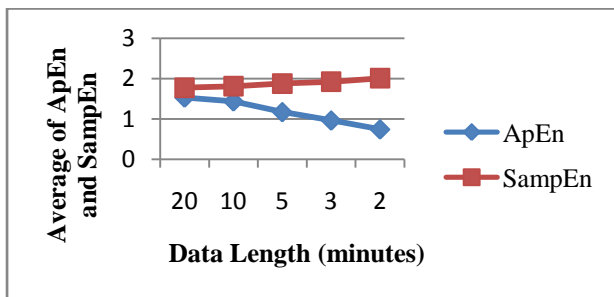


Fig 1. Comparison of Approximate and Sample Entropy

4. CONCLUSION

The results show that the length of RR intervals affects more on approximate entropy rather than sample entropy. While calculating approximate entropy, lower the value of ApEn higher will be the regularity. So, in case of ApEn the optimal RR interval data length is 20 minute. But, for Short-term HRV 5 minute data length is acceptable. Based on findings, it appears that SampEn is less sensitive to changes in data length and is more reliable for short data sets. So, it is possible to reduce the measuring time of sample entropy up to 2 minutes but in case of ApEn it is impossible to reduce the measuring time because of high value of Relative Error.

In the future there is possibility to analyze the different nonlinear parameters such as Poincare plot, Correlation, ApEn, SampEn and Shannon Entropy on 5 minute and 3 minute data length. Comparison should be done between 5 minute and 3 minute data length to show whether 3 minute data would differ significantly from the results of 5 minute data which is regularly used for the short term HRV analysis.

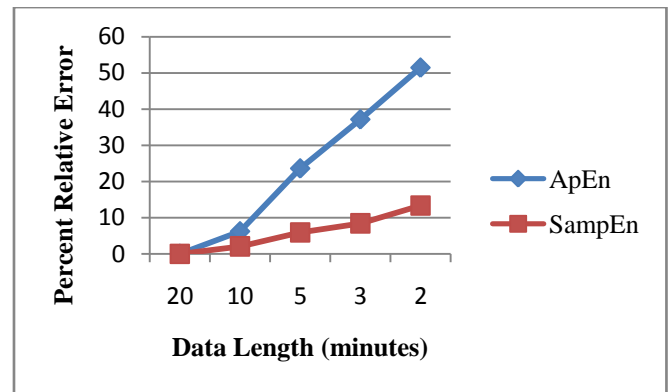


Fig 2. Relative Error of a.) Approximate Entropy and b.) Sample Entropy

5. REFERENCES

- [1] Afonso, V. X. 1993. ECG QRS detection. Biomedical Digital Signal Processing, chapter 12, 237-264. Prentice Hall, New Jersey.
- [2] Baselli, G., Cerutti, S., Civardi, S., Lombardi, F., Malliani, A., Merri, M., Pagani, M., and Rizzz, G. 1987. Heart rate variability signal processing: a quantitative approach as an aid to diagnosis in cardiovascular pathologies. International Journal of Bio-Medical Computing, vol. 20, 51-70.
- [3] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. 1996. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Circulation 93:1043-65.
- [4] Pichon, A., Roulaud, M., Antoine-Jonville, S., de Bisschop, C., Denjean. 2006. A Spectral analysis of heart rate variability: Interchangeability between autoregressive analysis and fast Fourier transform. Electrocardio 39:31-7.
- [5] Pincus, S. M. 1991. Approximate entropy as a measure of system complexity. In Proc. Nat. Acad. Sci. U.S.A., vol. 88, 2297-2301.
- [6] Haitham, M., Al-Angari* and Sahakian, A.V. 2007. Use of Sample Entropy Approach to Study Heart Rate Variability in Obstructive Sleep Apnea Syndrome. IEEE Transactions on Biomedical Engineering, vol. 54, 10.
- [7] Brennan, M., Palaniswami, M., and Kamen, P. 2001. Do Existing Measures of Poincaré Plot Geometry Reflect Nonlinear Features of Heart Rate Variability. IEEE Transactions on Biomedical Engineering, vol. 48, no. 11.
- [8] Lin, G. H., Chang, Y. H., and Lin, K. P. Comparison of Heart Rate Variability Measured by ECG in Different Signal Lengths," Journal of Medical and Biological Engineering, 25(2): 67-71
- [9] <http://www.physionet.org/physiobank/database/mitdb/> available at, MIT-BIH Arrhythmia Database.

- [10] <http://www.physionet.org/physiobank/annotations.shtml> available at, PhysioBank Annotations.
- [11] Pincus, S. M., and Huang, W. M. 1992. Approximate entropy: Statistical properties and applications. *Commun. Statist. Theory Meth*, vol. 21, 3061–3077.
- [12] Richman, J. S. and Moorman J. R. 2000. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.*, vol. 278, H2039–H2049.
- [13] Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-aho, P. O., and Karjalainen, P. A. 2014. Kubious HRV- Heart rate variability analysis software. *Comp Methods and Prog in Biomed.* 113:210-220.