Independence and symbolic independence of nonstationary heartbeat series during atrial fibrillation

Camillo Cammarota * Enrico Rogora †

December 22, 2004

Abstract

Heartbeat intervals during atrial fibrillation are commonly believed to form a series of almost independent variables. The series extracted from 24 hours Holter recordings show a non stationary behavior. Because of non stationarity it is difficult to give a quantitative measure of independence. In this paper we use and compare two methods for this. The first is a classical method which models a non stationary series using a linear Gaussian state space model. In this framework the independence is tested on the stationary sequence of the residuals. The second method codes data into permutations and tests the uniformity of their distribution. This test assumes as null hypothesis a weaker form of independence which we call symbolic independence. We discuss some advantages of symbolic independence in the context of heartbeat series. We analyze the time series of heartbeat intervals from 24 hours Holter recordings of 9 subjects with chronic atrial fibrillation and find that the detrended series is a zero or one memory process for 83% of regular segments and is symbolically independent for 93% of segments.

Keywords: Non stationary time series, state space model, symbolic dynamics, independence test, Holter analysis, atrial fibrillation.

*Dept. of Mathematics, University of Rome "La Sapienza", cammar@mat.uniroma1.it
†Dept. of Mathematics, University of Rome "La Sapienza", rogora@mat.uniroma1.it

1 Introduction

The heartbeat time series is the sequence of time intervals between two consecutive R peaks of the ECG, corresponding to the depolarization of the ventricles (systole). The series extracted from 24 hours Holter ECG consists of about 10^5 terms and is called RR sequence. Each beat is labeled by the analyzing software, which recognizes normal QRS complexes, ectopic ventricular beats, artifacts and so on.

The heartbeat time series in normal subjects is highly non stationary, due to the influences of the environmental stimuli on the heart, causing significant trends on different time scales. These trends superimpose to the fluctuations due to the autonomous heart dynamics.

Atrial fibrillation is an arrhythmia characterized by multiple re-entrant waveforms within the atria bombarding the atrioventricular node which becomes relatively refractive to conduction. This causes a totally irregular, often rapid, ventricular rate. The QRS complex has however the same appearance of a normal beat.

In atrial fibrillation fluctuations of RR sequence are larger than in normal case and their correlations are weaker. This has already been noted in the late sixties in [13] which showed that in stationary segments of RR sequences adjacent intervals are almost independent. The assumption of stationarity is no longer acceptable for segments of RR sequences which come from 24 hours Holter recordings, where non stationary behavior appears at both short time and long time scale. Examples of long time trends are circadian rhythms [14]. The question if atrial fibrillation is deterministic, eventually chaotic, or a purely stochastic phenomenon has been addressed in [20].

Heartbeat time series can be investigated using standard linear ARIMA models [5], and in normal subjects this gives autoregressive models of order $p \geq 16$ [4]. In atrial fibrillation the situation is different: when the series is detrended, the residuals show a low autoregressive order ($p \leq 1$); furthermore the series shows an heteroschedastic behavior: the greater is the mean level, the greater is the variance, see [21] and [8].

In this paper we are interested in studying the statistical properties of the Holter recorded RR sequences during atrial fibrillation. At this aim we consider of primary importance two problems: suitably estimating the trend and filtering out the beats for which the QRS complex is not labeled as normal, like ectopic ventricular beats and artifacts.

As to the first problem, a standard approach for detrending a non stationary time series is to use linear Gaussian state space models, see [11]. In these models the trend is extracted by means of a Kalman filter; the residuals, obtained by subtracting the trend, form an i.i.d. Gaussian sequence and

the model can be checked by performing independence and normality tests on the residuals. Usually state space methods are applied for time series with few hundreds of elements. Our RR series, which are extracted from 24 hours Holter recordings, are much longer, hence for applying this method we divide the series in 1000 beats segments. For each segment we fit the residuals with an autoregressive model of order zero or one. This allows us to make a precise quantitative statement about the independence of residuals. Another method for extracting a trend which has received much attention in the analysis of cardiac signals is detrended fluctuation analysis [17]. However we found easier to frame our results using the more versatile state space models.

A further possibility for analyzing non stationary series is to use a symbolic dynamic approach, which has recently received much attention, in particular for physiological signals. Various methods have been adopted, based on coding short strings into words or permutations ([22], [18], [3], [16], [2], [12], [9]). In particular atrial fibrillation has been investigated with these methods in [1], [7], [9].

The coding into permutations provides a very natural method for testing independence, based on a standard property: if the series is an i.i.d. sequence, the distribution of permutations is always uniform. Since the coding reduces information, the notion of independence for coded series, which we call *symbolic independence*, is weaker than independence of the original one.

A second problem is that of filtering out non normal beats, since we want to separate the beats due to atrial fibrillation from those not fired by electrical activity of the atria, like ventricular ectopic beats. In order to do this we label all non normal beats as NA (missing value). Then we analyze the series either by filling the missing beats by using Kalman filter or by performing symbolic analysis only on segments without missing beats. We believe that filling the data by using Kalman filter is something quite artificial and not appropriate in presence of many missing values. We have chosen to compare the two methods on segments of 1000 beats in which at least 90% of the beats were normal. We have found that the results on independence obtained with Kalman filter agree with those obtained with symbolic analysis. This separation of beats into classes according to their labels is also crucial for the validation of mathematical models of atrioventricular node conduction during atrial fibrillation, as the one proposed by Zeng and Glass in [23] (see also [15] and [10]). In fact these models refer only to beats for which the QRS complex is normal.

We explain our use of state space models in sections 2 and 3, and permutation coding in section 4. In section 5 we compare the two methods analyzing 24 hours Holter recordings of 9 subjects with atrial fibrillation and

state our final remarks in section 6.

2 The state space model

We review some basic facts about autoregressive models following [5] and about state space models following [11] to which we refer for an extensive treatment.

A basic model for stationary time series is the autoregressive model of order p, AR(p), in which data are modeled as a jointly Gaussian random vector $x = (x_1, \ldots, x_N)$; the dependence is defined by

$$x_i = \sum_{j=1}^{p} \phi_j x_{i-j} + w_i, \qquad w_i \sim N(0, \sigma_w^2)$$

where the ϕ_j's are coefficients, and the w_i's are a sequence of independent Gaussian variables. In our analysis we shall use only the AR(1) model

$$x_i = \phi_1 x_{i-1} + w_i \tag{1}$$

and the trivial AR(0) model $x_i = w_i$.

A measure of independence is given by the autocorrelation function which is defined by

$$\rho_k = \mathrm{Cov}(x_i, x_{i-k})$$

For AR(1) models one has $\rho_k = \phi_1^k$.

In case of non stationary series a basic approach is provided by the state space models. A state space model is based on two jointly Gaussian random vectors $x = (x_1, \ldots, x_N)$ and $\alpha = (\alpha_1, \ldots, \alpha_N)$ such that

$$x_i = \alpha_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

$$\alpha_{i+1} = \alpha_i + \eta_i, \quad \eta_i \sim N(0, \sigma_\eta^2)$$

for $i = 1, \ldots, N$. Here the sequences of normal random variables η_i's and ϵ_i's are mutually independent and independent on the α_i's.

The logarithm of the joint probability density of x and α is, apart from constants,

$$-\frac{1}{2\sigma_\eta^2} \sum_{i=1}^{N-1} (\alpha_{i+1} - \alpha_i)^2 - \frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^{N} (x_i - \alpha_i)^2$$

The vector α is the *local level* (or state vector) and the vector x is the vector of observations. In this model it is possible to estimate the distribution

of α given the observations x. Since the conditional probabilities of the α given x are Gaussian, this reduces to estimate the means and the variances of the α_i's. In particular in the filtering approach one defines

$$a_i = E(\alpha_i | x_1, \ldots, x_{i-1}), \quad i = 2, \ldots, n \text{ and } a_1 = x_1$$

The computations of the means a_i and of the related variances give rise to recursive relations, which constitute the Kalman filter. In particular one has the relations

$$a_{i+1} = a_i + K_i(x_i - a_i), \quad a_1 = x_1, \quad i = 2, \ldots, N \tag{2}$$

where the sequence K_i depends only on the variances σ_ϵ^2 and σ_η^2. This sequence rapidly converges to a value K. In the hypothesis that

$$h = (\sigma_\eta / \sigma_\epsilon)^2 << 1$$

neglecting the terms of order greater or equal to h, one has

$$K \simeq \sigma_\eta / \sigma_\epsilon \tag{3}$$

These variances can be estimated from the data by means of a maximum likelihood method. The variance of the residuals $v_i = x_i - a_i$, denoted F_i, also converges to a limit, which in the above approximation is $F \simeq \sigma_\epsilon^2(1 + \sigma_\eta/\sigma_\epsilon)$. The Kalman filter algorithm is implemented in the package `StructTS` of the free statistical software `R` [24] that we have used in our data analysis. The main features of this model are: it provides an estimate of the trend, defined by the sequence a_i; the residuals v_i are mutually independent Gaussian variables. Hence the model diagnostic consists essentially of an independence test and a normality test on the residuals. An additional criterion to evaluate the suitability of the model, at least in our setting, is $h << 1$. In this case the level sequence a_i is weakly influenced by observations, and can be indeed interpreted as a regular trend by equation (2).

3 Estimation of trend

Our previous investigations suggest that the heartbeat time series in atrial fibrillation can be modeled as a non stationary sequence of independent random variables [8], [9]. In time series analysis ARIMA models are often used [5], and for removing non stationarity one has to differentiate the series. In heartbeat series this is a very efficient method, in particular for normal subjects. Actually the autocorrelation function, which for the original series has

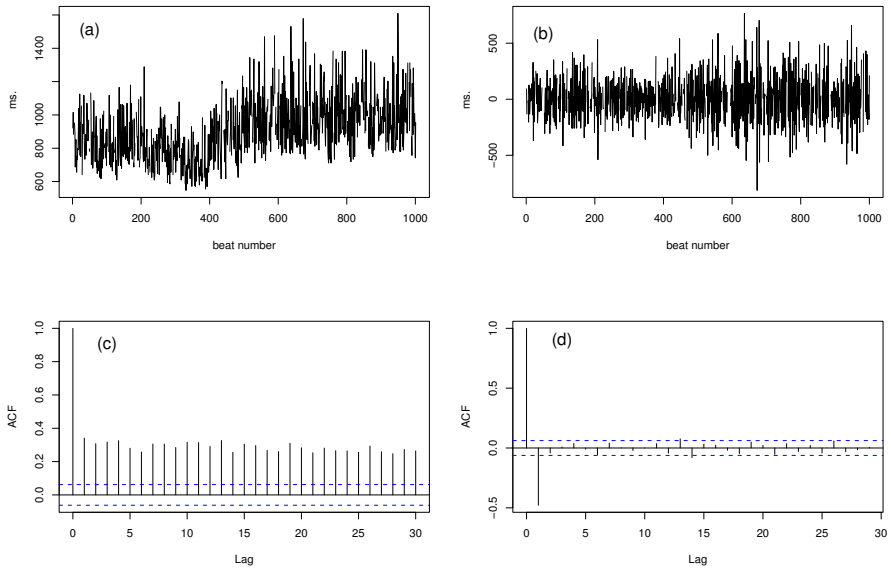

Figure 1: (a): Plot of non stationary segment of length $N = 1000$ of the time series; (b): Plot of the differences of segment (a); (c): Autocorrelation of the segment (a) ; (d) Autocorrelation of the segment (b). The values of the autocorrelation within the dotted lines are considered non significant.

a very slow decay, after differentiation has a much faster decay [6]. In atrial fibrillation series after differentiation the autocorrelation function has only one significant value at lag 1, which is a little above -0.5, as shown in figure 1. One can easily show that differentiating a sequence of i.i.d. variables, the autocorrelation is zero for lags greater or equal than 2 and takes exactly the value -0.5 for lag 1. Hence the data in atrial fibrillation should be modeled by a non stationary sequence of weakly dependent variables, and one has to use a more subtle method for removing non stationarity than differentiation.

One of the features of the inter beat intervals in atrial fibrillation is a non stationarity of the variance. More precisely the variance is increasing with respect to the local value of the mean ([21], [8]). According to a common method used to stabilize the variance, we take the logarithm of the data. In order to estimate the mean level we use the Kalman filter defined in section 2. The mean level is shown as a solid line in figure 2(a). The result of the filter algorithm provides also the two variances σ_η^2 and σ_ϵ^2. In figure 2 $h = 0.005$, hence $K \simeq 0.07$. The small value of K implies that the level sequence a_i depends weakly on the data values according to equation (2), so

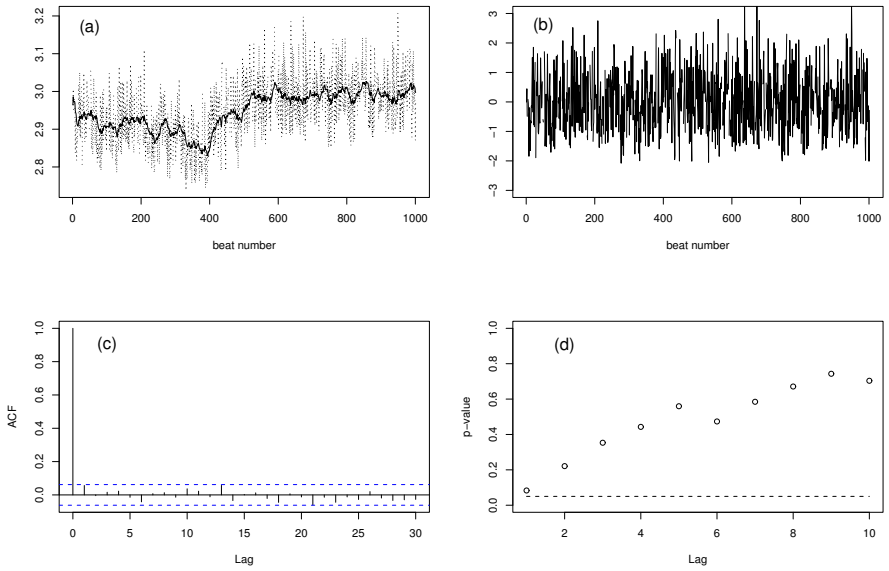

Figure 2: (a): Logarithmic values (base 10) of the series in figure 1(a) (dotted) and the level line (solid); ratio of the variances $\sigma_\eta^2/\sigma_\epsilon^2 = 0.005$; (b): Standardized residuals; (c): Autocorrelation of residuals; values within the dotted lines are non significant; (d) Independence test (Ljung-Box) for the residuals up to the lag 10, and significance level 0.05. Values under the threshold (dotted line) reject independence hypothesis.

it appears much more stable than the series. For this reason one can assume this sequence as an estimation of the trend.

The Kalman filter algorithm provides a simple method for treating both the mean level and its variance also when there are missing values. In particular the mean level sequence a_i is defined to be constant in correspondence of missing data values; in other words the same recursion equations apply with $K_i = 0$ for those i's for which x_i is missing; the residuals are defined to be zero. The plot of data in figure 1 contains missing values, although this is not evident in the picture.

In the assumptions of the state space model the residuals with respect to the Kalman filter form a sequence of normal independent variables. Hence the diagnostic of the model consists of a test of independence and a test of normality. The autocorrelation function provides a first insight on the independence. Under this assumption the values of the autocorrelation function are normally distributed and if N is the length of the sequence a 95% confidence interval is $(-1.96/\sqrt{N}, 1.96/\sqrt{N})$. In figure 2(c) the values are inside

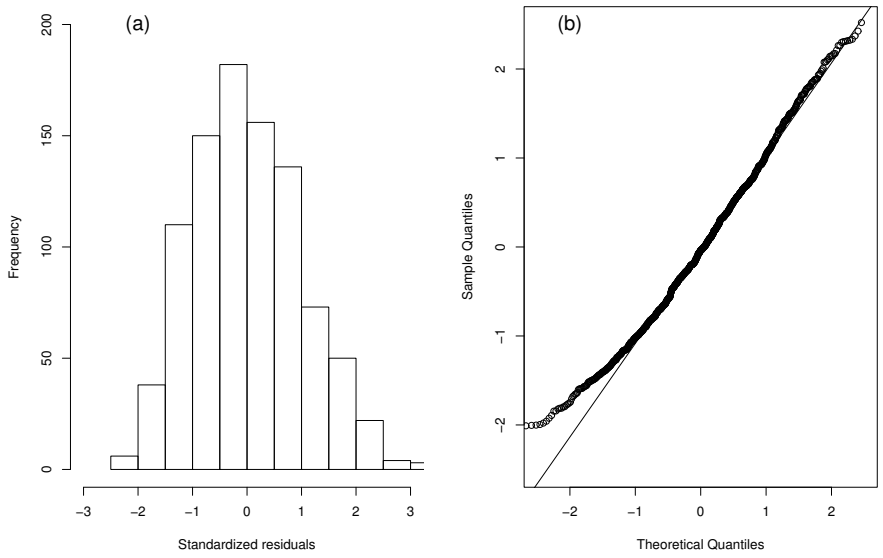

Figure 3: Normality check for residuals of figure 2(b). (a): Histogram of standardized residuals; (b): qq-plot of normal standard quantiles versus quantiles of standardized residuals.

this interval. A standard test of independence for residuals is the Ljung-Box test [5]. This test checks the hypothesis that the first k values of the autocorrelation are compatible with the independence assumption. An usual choice is $k = 10$. We reject the hypothesis if any of the p-values is less than 0.05. As to normality we show in figure 3 an example of histogram of residuals and of qq-plot.

4 Symbolic dynamics

In this section we follow a method we have developed in [9]. Here we recall the essential points in order to get a self contained exposition. We code short segments of fixed length n into permutations in the following way. In $\mathbb{R}^n$ let us consider the subset Δ (called *big diagonal*) which is the set of points $(x_1, \ldots, x_n)$ for which there exist at least two indexes i , j such that $x_i = x_j$ and denote $\mathbb{R}^n \setminus \Delta$ by $\mathbb{R}^n_{\neq}$. Let S_n denote the symmetric group, i.e. the set of permutations of $\{1, \ldots, n\}$. We use one line notation for permutations, i.e.

$$(i_1, i_2, \ldots, i_n)$$

denotes the permutation

$$\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

We define a function $\Pi : \mathbb{R}^n_{\neq} \to S_n$ in the following way:

$$\Pi(x_1, \dots, x_n) = (\pi(1), \dots, \pi(n))$$

where

$$\pi(i) = 1 + \#\{j : x_j < x_i\}$$

(The symbol # denotes cardinality). Note that with this definition, if π^{-1} is the inverse of π, then

$$x_{\pi^{-1}(1)} < x_{\pi^{-1}(2)} < \cdots < x_{\pi^{-1}(n)}$$

For example, if $(x_1, \dots, x_4) = (181, 32, 42, 115)$, then $\pi = (4, 1, 2, 3)$, $\pi^{-1} = (2, 3, 4, 1)$ and

$$x_2 < x_3 < x_4 < x_1.$$

Let us denote by $(x_1, \dots, x_n)$ a collection of n independent random variables identically distributed with probability density f positive and absolutely continuous with respect to the Lebesgue measure of $\mathbb{R}$, and let P be the product probability measure on $\mathbb{R}^n$. The subsets

$$A_\pi = \{(x_1, \dots, x_n) \in \mathbb{R}^n_{\neq} : \Pi(x_1, \dots, x_n) = \pi\}$$

parameterized by the permutations $\pi \in S_n$ give a partition of $\mathbb{R}^n_{\neq}$. We define the probability measure P_Π over S_n as

$$P_\Pi(\pi) = P(A_\pi)$$

The probability P_Π is uniform on S_n, i.e.

$$P_\Pi(\pi) = 1/n!$$

(see [19]). We note that this property does not depend on the density f. We say that a sequence of random variables is *n-symbolically independent* if the distribution P_Π is uniform.

Let $x_1, \dots, x_N$ be a i.i.d. random sequence and let $\sigma \in S_n$. We define the estimator

$$T(\sigma) = \left[\frac{N}{n}\right]^{-1} \sum_{i=1,\ i \equiv 1 \bmod (n)}^{N-n+1} \chi(\Pi(x_i, \dots, x_{i+n-1}) = \sigma) \tag{4}$$

where: χ is the indicator function which takes the value 1 when the equality $\Pi(x_i, \ldots, x_{i+n-1}) = \sigma$ is true and 0 otherwise; square brackets denote integer part, hence $\left[\frac{N}{n}\right]$ is the number of consecutive disjoint intervals of length n in which we can split the series; the indexes in the sum are the minimum of these intervals. Let I be the set of these indexes. For brevity we shall write

$$T(\sigma) = \left[\frac{N}{n}\right]^{-1} \sum_{i \in I} \chi_i(\sigma).$$

We notice that the random variables $\chi_i(\sigma)$ for $i \in I$ are independent. Hence the $n!$ variables $T(\sigma)$ form a multinomial vector with $\sum_{\sigma \in S_n} T(\sigma) = \left[\frac{N}{n}\right]$ and with probabilities $P_\Pi(\sigma) = \frac{1}{n!}$.

The statistics

$$\left[\frac{N}{n}\right] \sum_{\sigma \in S_n} \frac{(T(\sigma) - \frac{1}{n!})^2}{\frac{1}{n!}} \tag{5}$$

is asymptotically distributed as a $\chi^2(n! - 1)$ (as $N \to \infty$) and we use the standard χ^2 test for goodness of fit (see [19]). We refer to this test for symbolic independence as to *permutation test* (PT).

5 Data analysis

We have analyzed the RR time series of 24 hours Holter recordings of 9 cases of atrial fibrillation provided by the Department of Cardiology of our University. The data were recorded using an Holter equipment with sampling frequency of 180Hz (Rozinn Electronics, Glendale, USA); the analyzer software labels each beat with a code number. We have considered only beats which are labeled normal and we have coded NA (Not Available) all other beats.

We divide each time series into segments of length 1000 and give a code *NA* to those segments in which the number of NA beats is greater than 100; these segments are discarded. For the remaining ones we start our analysis by fitting a state space model in each segment. After detrending, the residuals are computed. If the ratio of variances $(\sigma_\eta/\sigma_\epsilon)^2$ is greater than 0.01 we label the segment as *NS* (Non Smooth). Otherwise an LB test is performed for the first $k = 10$ lags. If all the p values are greater than 0.05, then we do not reject the independence hypothesis and attribute a code *AR0* to the segment. If not, we fit the sequence v_i of residuals with an $AR1$ model

$$v_i = \phi_1 v_{i-1} + w_i \tag{6}$$

N	SEG	NA	NS	AR0	AR1	NAR	PT	AR0&PT
1	107	9	0	73	14	11	93	72
2	99	14	0	31	43	11	73	29
3	157	14	4	53	35	51	137	51
4	106	0	9	71	11	15	102	69
5	110	5	12	38	42	13	96	34
6	107	65	1	16	20	5	36	14
7	92	38	1	21	28	4	52	21
8	100	44	0	39	10	7	50	35
9	116	20	0	67	16	13	91	66

Table 1: The table summarize the analysis of 9 RR sequences of atrial fibrillation divided into segments of length 1000. Column N: case number; column SEG: number of segments analyzed for each case; column NA: number of segments with more than 100 NA; column NS: number of segments with no more than 100 NA with non smooth trend; column $AR0$: number of segments with no more than 100 NA and smooth trend whose residuals, after detrending, are independent; column $AR1$: number of segments with no more than 100 NA and smooth trend whose residuals, after detrending, fit an AR1 model; column NAR: number of segments with no more than 100 NA and smooth trend whose residuals, after detrending, are neither $AR0$ nor $AR1$; column PT: number of segments with no more than 100 NA which pass the permutation test; column $AR0\&PT$: number of segments labeled $AR0$ and PT

where w_i is a sequence of normal independent variables (see [5]). We use the package `ARIMA` implemented in `R` for fitting the model and perform LB test on residuals: if the model is not rejected we code the segment as $AR1$. Otherwise, we code the segment as NAR. We note that the coefficient ϕ_1 has typical values positive and less than 0.1.

We then perform the permutation test described in section 4 to all segments which are not coded as NA. This test depends on one parameter, i.e. the length of the permutations which are considered. In this work this parameter is chosen to be 3. We code each segment as PT if symbolic independence is not rejected by permutation test at the 0.05 significance level. Finally we compute the number of segments that are coded both $AR0$ and PT, i.e. the segments that are independent in both tests, denoted as $AR0\&PT$.

The results of this analysis are reported in table 1.

We give here some more explanation of the content of the table.

Column SEG: The variability of the number of segments in the 24 hours

Holter recording reflects the variability of heart rate among individuals.

Column *NA*: The number of *NA* segments may be very large (case 6 and 8) as well as very small (case 5) or even zero (case 4).

Column *NS*: The values are small and reveal that smooth trend extraction algorithm provided by Kalman filter performs well for atrial fibrillation. We shall call a smooth trend segment also a *regular* segment. Their number is $SEG - NA - NS$.

Column *AR0*: The values are extremely variable. In order to make comparison between different cases easier we define the following index

$$AR0n = \frac{AR0}{SEG - NA - NS}$$

This index is reported in table 2. Its average value is 0.54.

Column *AR1*: As above we define the index

$$AR1n = \frac{AR1}{SEG - NA - NS}$$

This index is reported in table 2. Its average value is 0.31.

Column *NAR*: This column is reported for completeness. Its values are obtained as $SEG - NA - NS - AR0 - AR1$

Column *PT*: As for *AR0* we define the index

$$PTn = \frac{PT}{SEG - NA}$$

whose values are reported in table 2. We remark that the values are very large (minimum is 86%), less variable than *AR0n* and *AR1n*; their average is 92%.

Column *AR0&PT*: One expects that the set of segments coded *AR0* is included in the set coded *PT* (i.e. *AR0*=*AR0&PT*), apart from differences due to accepting and rejecting errors of probabilistic tests. The observed little difference in the results reported in the table are compatible with this discrepancy.

We remark that permutation test has been applied to all "non *NA* segments" while the other tests has been applied only to *regular* segments.

We finally remark that the independence test of residuals checks correlations up to lag 10, while we have used a permutation test based on permutations of 3 elements only.

6 Conclusions

State space models provide a mathematical framework under which a suitable notion of independence can be tested for non stationary heartbeats during

N	AR0n	AR1n	PTn
1	0.75	0.14	0.95
2	0.37	0.51	0.86
3	0.38	0.25	0.96
4	0.73	0.11	0.96
5	0.41	0.45	0.91
6	0.39	0.49	0.86
7	0.40	0.53	0.96
8	0.70	0.18	0.89
9	0.70	0.17	0.95

Table 2: Comparison between indexes of independence computed from standard and permutation tests. Column N: case number; column $AR0n$: ratio of $AR0$ segments on segments which are neither NA nor NS; column $AR1n$: ratio of $AR1$ segments on segments which are neither NA nor NS; column PTn: ratio of PT segments on segments which are not NA.

atrial fibrillation. One of the advantages of this model is that we can give a meaningful interpretation of the elements of the model: the level sequence reflects the influence of the external world on the heart; the residuals reflect the autonomous dynamics of the heart. In our database of atrial fibrillation the autonomous dynamics can be modeled as an independent sequence for 54% of regular segments and as a first order autoregressive model for other 29%; in summary, the autonomous dynamics is a zero or one memory process for 83% of the regular segments.

On the other hand the autonomous dynamics is symbolically independent for 93% of non NA segments. Moreover, the symbolic dynamics approach is coherent with the state space one. More precisely, the idea that symbolic independence is weaker than the independence of residuals over $AR0$ segments has been confirmed by our data analysis.

The symbolic approach has some advantages over state space models. First, it can be used to analyze time series with many missing values. Second, permutation test is not dependent on the distribution of the original series. Third, it can be applied also to smooth trend data without detrending. Finally, it provides a possible method for measuring a distance from independence. This can be used, for instance, for distinguishing the heart beat series of atrial fibrillation from normal ones, where the smooth trend condition does not hold; see [9]. On the other hand it has at least one disadvantage since coding a series into permutations obviously causes a loss of information.

This statistical analysis can also be used to validate models for nodal conduction in atrial fibrillation like [23]. Actually the RR sequences simulated using this model, according with some preliminary work, fit AR0 or AR1 models.

Our results may have clinical applications too. A large percentage of independence can be interpreted as a loss of efficiency of the control mechanism of autonomous dynamics of the heart.

Acknowledgments We thank Giuseppe Germanò and Maria Ambrosini of Policlinico of our University for providing data of atrial fibrillation and Giuseppe Guarini for useful discussions.

References

[1] Albert, C.-C. Yang, Shu-Shya Hseu, Huey-Wen, Yien, Ary L. Goldberger, C.-K. Peng "Linguistic Analysis of the Human Heartbeat Using Frequency and Rank Order Statistics", *Phys. Rev. Lett.*, **90** (2003) , 108103-1,108103-4.

[2] Ambrosini M., Cammarota C., Guarini G. "Heart beat stationarity in heart transplanted patients", *Clin. Ter.*, **152** (2001), 363-366

[3] Bandt C., Pompe B. "Permutation Entropy: A natural Complexity Measure for Time Series", *Phys. Rev. Lett.* **88** (2002) 174102-1:174102-4

[4] Boardman A., Schlindwein F. S., Rocha A. P., Leite A. "A study of the optimum order of autoregressive models for heart rate variability". *Physiological measurements* **23** (2002) 325-336.

[5] Brockwell P.J., Davis R.A. *Time Series: Theory and Methods* New York, Springer, 1987.

[6] Cammarota C. "Estimating correlation exponents of the heartbeat time series" *Int. J. Bifurcation and Chaos*, **10** (2000) , 1513-1520

[7] Cammarota C., Guarini G., Ambrosini M. "Analysis of stationary periods of heart rate via symbolic dynamics", *Medical Data Analysis*, Lecture Notes in Computer Science vol. 2526, Colosimo, Giuliani and Sirabella eds. Springer 2002

[8] Cammarota C., Guarini G., Rogora E., Ambrosini M. "Non stationary model of the heart beat time series in atrial fibrillation", *Mathematical modeling and computing in Biology and Medicine, Proceedings of 5th ESMTB Conference, Milan 2002* Ed. V. Capasso, Milano, Società Editrice Esculapio 2003. Available on line at the web page `http://www.mat.uniroma1.it/~cammar`

[9] Cammarota C., Rogora E."Estimating independence in time series via universal distributions of permutations and words" Accepted for publication on *Int. J. Bifurcation and Chaos*. Preprint n. 6/2004 of the Department of Mathematics, University "La Sapienza" of Rome. Available on line at the web page `http://www.mat.uniroma1.it/~cammar`

[10] Cammarota C., Rogora E. *Validation of nonstationary model of atrial ventricular node conduction in atrial fibrillation* submitted to Proceedings of SIMAI congress 2004, Venice and available on line at the web page `http://www.mat.uniroma1.it/~cammar`

[11] Durbin J., Koopman S. J., *Time series analysis by state space methods*, Oxford, Oxford University Press, 2001.

[12] Germanò G., Piccirillo G., Cammarota C., Guarini G., Rogora E., Cacciafesta M. "A new measure of acceleration of heart rate: dependence on age and comparison with time domain conventional heart rate variability measures", Preprint n. 42/2003 of the Department of Mathematics, University "La Sapienza" of Rome. Available on line at the personal web page `http://www.mat.uniroma1.it/~cammar`

[13] Goldstein R. E., Barnett G. O. *A statistical study of the ventricular irregularity of atrial fibrillation* Computers and biomedical research **1**, 146-161 (1967)

[14] J. Hayano, S. Sakata, A. Okada, S. Mukai, T. Fujinami, "Circadian Rhythms of Atrioventricular Conduction Properties in Chronic Atrial Fibrillation With and Without Heart Failure", *J. Am. Coll. Card.* **31**, 158-166 (1998).

[15] Jorgensen P., Schafer C., Guerra P. G., Talajic M., Nattel S., Glass L. *A Mathematical Model of Human Atrioventricular Nodal Function Incorporating Concealed Conduction* Bull. Math. Biol. **64** 1083-1099, (2002)

[16] Keller K., Lauffer H., "Symbolic analysis of high dimensional time series", *Int. J. Bifurcation and Chaos* **13** 2657-2668 (2003)

[17] C.K.Peng, S.Havlin, H.E. Stanley, A.L.Goldberger, Quantification of scaling exponents and crossover phenomena in non stationary heartbeat time series, *Chaos* **5**, 82-87 (1995)

[18] Porta A., Guzzetti S., Montano N., Furlan R., Pagani M., Malliani A., Cerutti S., *Entropy, Entropy Rate, and Pattern Classification as Tools to Typify Complexity in Short Heart Period Variability Series* IEEE Trans. Biom. Eng. **48** 1282-91 (2001)

[19] Rohatgi V. K., *An introduction to Probability Theory and Mathematical Statistics*, New York, Wiley ed., 1976

[20] K. M. Stein, J. Walden, N. Lippman, B. B. Lerman, "Ventricular Response in Atrial Fibrillation: random or deterministic", *Am. J. Phisiol.*, **277**, H452-H458 (1999).

[21] Tateno K., Glass L. *Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and deltaRR intervals* Med. Biol. Eng. Comput., **39** 664-71, (2001)

[22] Zebrowski J.J., Poplawska W., Baranowski R., Buchner T. "Symbolic dynamics and complexity in a physiological time series" *Chaos, solitons and Fractals* **11** 1061-1075 (2000)

[23] Zeng W., Glass L. "Statistical properties of heart beat intervals during atrial fibrillation" *Physical Review E* vol. **54** 1779-1784 (1996)

[24] `http://www.R-project.org`