

The difference-sign runs length distribution in testing for serial independence

Camillo Cammarota

*Dept. of Mathematics, University of Rome "La Sapienza",
P.le A. Moro 2, 00185, Rome, Italy*

Abstract

We investigate the sequence of difference-sign runs lengths of a time series in the context of non parametric tests for serial independence. This sequence is under suitable conditioning a stationary sequence and we prove that the normalized correlation of two consecutive runs lengths is small (≈ 0.0427). We use this result in a test based on the relative entropy of the empirical distribution of the runs lengths. We investigate the performance of the test in simulated series and test serial independence of cardiac data series in atrial fibrillation.

Key words: runs test, difference-sign, serial independence, entropy, time series, atrial fibrillation

2000 MSC: 62G10, 62M10

1. Introduction

The difference-sign runs test plays a central role in non parametric tests for serial independence. This test is often used in time series analysis for testing the independence of the residuals in fitted deterministic trend models [Brockwell and Davis, 2002]. The definition of difference-sign runs can be given in terms of occurrence times of extrema (maxima and minima) in the series. Let X_1, X_2, \dots be a sequence of continuous random variables (r.v.), whose indexes are referred to as "times". Denote $A = \{x_1 < x_2, x_2 > x_3\} \subset R^3$ the event corresponding to a maximum, $B = \{x_1 > x_2, x_2 < x_3\} \subset R^3$ the event corresponding to minimum and $E = A \cup B$ corresponding to an extreme. For instance we say that in i a maximum occurs if $(X_{i-1}, X_i, X_{i+1}) \in A$.

Let T_1, T_2, \dots denote the occurrence times of the event E , with $T_1 \geq 2$ and $T_i < T_{i+1}$ and let H_i denote the associated recurrence times

$$H_i = T_{i+1} - T_i \quad i = 1, 2, \dots$$

Email addresses: cammar@mat.uniroma1.it (Camillo Cammarota),
<http://www.mat.uniroma1.it/people/cammarota/> (Camillo Cammarota)

The intervals defined by the T_i are the difference-sign runs, hereafter called simply “runs”, and the H_i are their lengths. The classical runs test is based on the statistic defined as the number of runs. Since this statistic is based on ranking of the values, it gives rise to a distribution-free test [Bradley, 1968]. In a sequence of independent and identically distributed (i.i.d.) X_1, \dots, X_n continuous r.v. of size n the runs statistic, denoted N , is known to be approximated by a normal of mean $(2n - 1)/3$ and variance $(16n - 29)/90$ [Levene and Wolfowitz, 1944, Wolfowitz, 1944]. For practical implementation of this kind of tests we also refer to Sachs [1982]; for different definitions of runs we refer to Balakrishnan and Koutras [2001]. The covariance matrix of the vector N_1, \dots, N_p , where N_j is number of runs of length j is also known [Levene and Wolfowitz, 1944]. These results have been used to test pseudo random numbers generators [Knuth, 1988].

Physiologic signals provide a lot of data series in which one has to use robust methods such as non parametric tests for serial independence. There are at least two reasons: the data distribution is generally non normal or not known and outliers are often present. For a public database and bibliography we refer to Goldberger et al. [2000]. We consider here the time series of RR intervals (see below) extracted from the electrocardiograms (ECG). The amount of randomness in these series has been widely investigated in physical literature related to various normal and pathological conditions(see for instance Urbanowicz et al. [2007] and references therein). In some pathologies, such as atrial fibrillation, the RR series is characterized by a high level of uncorrelated noise.

However, at our knowledge, is still lacking an analysis based on rigorous testing procedures. A first step in this direction is in Cammarota and Rogora [2005] and suggests that this series can be modeled by a smooth deterministic trend plus an autoregressive model of order 0 or 1.

The use of runs test in these data has, in addition to the motivations above mentioned, a third one: the runs have a direct physiologic interpretation in terms of acceleration and decelerations of heart rate (see below). Hence a runs test based on the distribution of the runs lengths seemed to be more appropriate than the standard one based on the number of runs. Unfortunately the sequence of runs lengths extracted from an i.i.d. sequence are not independent. The joint probability of two consecutive runs lengths was computed in Levene and Wolfowitz [1944]. This lack of independence has been noticed later in Knuth [1988] and Chen and Kelton [2003], where a modified definition of runs was adopted to assure independence.

One of the results of the present paper concerns the dependence structure of the sequence of runs lengths H_1, H_2, \dots . Under suitable conditioning this sequence is stationary [Cammarota and Curione, 2008] and ergodic. In this setting we compute the lag one normalized autocorrelation of this sequence, and show that it is positive, but very small (≈ 0.0427). This implies that a sample H_1, \dots, H_n of runs lengths extracted from an i.i.d. sequence can be considered approximately as a random sample. This leads us to a test for serial independence based on the distribution of the runs lengths. The statistic is defined as the relative entropy of the empirical runs distribution with respect to

the runs distribution of i.i.d.. The performance of the test is compared to the one of the standard runs test both in simulated processes and in data series of atrial fibrillation. An example is given in which the entropy test outperforms the standard one in detecting dependence.

2. The correlation of two consecutive runs lengths

The main new result in this section contained in Proposition 4 and eq. (11) states that the correlation of two consecutive runs lengths extracted from an i.i.d. sequence is small (≈ 0.0427); the exact value is given by eq.(11). The proof requires some preliminary results (Propositions 2 and 3), which go back to Levene and Wolfowitz [1944], whose proofs are in the Appendix. Some basic results on the sequence of runs lengths are in Proposition 1 [Cammarota and Curione, 2008].

The first result concerns the conditions under which the sequence H_1, H_2, \dots of runs lengths extracted from an i.i.d. sequence is stationary. For this we need to introduce the sequence

$$X'_i = (X_i, X_{i+1}, X_{i+2}) \in R^3 \quad i = 1, 2, \dots$$

and consider the probability conditioned to the event $\{X'_1 \in E\}$, i.e. ‘in 2 there is an extremum’. Following general properties for stationary and ergodic sequences of recurrence times [Breiman, 1993], the sequence H_1, H_2, \dots is stationary if the probability is conditioned to start in an extreme. A second basic result is the well known distribution free equation [Rohatgi, 1976]

$$P(X_1 < \dots < X_s) = \frac{1}{s!} \tag{1}$$

where X_i are i.i.d. variables with continuous product distribution P . We summarize the basic properties of runs lengths in the following proposition.

Proposition 1 If X_i are i.i.d. variables with continuous distribution, the sequence of runs lengths H_i is stationary and ergodic in the measure conditioned to the event $X'_1 \in E$. The conditional mean is

$$\mathbb{E}(H_1 | X'_1 \in E) = \frac{3}{2} \tag{2}$$

Proof As a preliminary step we compute the probability of an extrema. By symmetry

$$P(X'_1 \in E) = P(X'_1 \in A) + P(X'_1 \in B) = 2P(X'_1 \in A)$$

and using eq.(1) one has

$$\begin{aligned} P(X'_1 \in A) &= P(X_1 < X_2, X_2 > X_3) \\ &= P(X_1 < X_2) - P(X_1 < X_2 < X_3) = \frac{1}{2} - \frac{1}{6} = \frac{1}{3} \end{aligned}$$

and

$$P(X'_1 \in E) = \frac{2}{3} \quad (3)$$

The sequence X'_i is stationary and from eq.(3) we get $P(X'_i \in E \text{ i.o.}) = 1$. The times H_1, H_2, \dots are finite a.e., they form a stationary sequence under the conditional probability $P(\cdot | X'_1 \in E)$. The stationary sequence X'_i is 2-dependent i.e. X'_i and X'_j are independent if $|i - j| > 2$ and not independent elsewhere. Hence it is also ergodic and hence H_i is an ergodic (not independent) sequence in the probability conditioned to $X'_1 \in E$. Eq. (2) can be obtained easily from the well known Kac formula for the expectation of the first recurrence time

$$\mathbb{E}(H_1 | X'_1 \in E) = \frac{1}{P(X'_1 \in E)} \quad (4)$$

□

Remarks By symmetry the probability of events generated by H_1, H_2, \dots conditioned to E is the same as conditioned to A or to B . For instance $P(H_1 = s | X'_1 \in E) = P(H_1 = s | X'_1 \in A)$. The mean $3/2$ is to be put into correspondence to the leading term $2/3$ of the expectation of the number N of runs in a series of n elements [Levene and Wolfowitz, 1944].

The univariate distribution and the first two moments are summarized in the following proposition, where we denote the conditional probability shortly as $P(\cdot | E)$.

Proposition 2 The conditional distribution of the runs lengths is

$$P(H_1 = s | E) = 3 \left[\frac{1}{s+2} \frac{1}{s!} - \frac{1}{s+3} \frac{1}{(s+1)!} \right], \quad s \geq 1 \quad (5)$$

and

$$\mathbb{E}(H_1 | E) = \frac{3}{2} \quad (6)$$

$$\mathbb{E}(H_1^2 | E) = 3(2e - \frac{9}{2}) \quad (7)$$

$$\text{Var}(H_1 | E) = 3(2e - \frac{21}{4}) \approx 0.560 \quad (8)$$

Proof: See Appendix

Proposition 3 The joint distribution of two consecutive runs lengths is

$$P(H_1 = s, H_2 = t | E) = 3 \left[\frac{1}{s+t+1} \frac{1}{s!} \frac{1}{t!} - \frac{1}{s+t+2} \frac{1}{s!} \frac{1}{(t+1)!} \right. \\ \left. - \frac{1}{s+t+2} \frac{1}{(s+1)!} \frac{1}{t!} + \frac{1}{s+t+3} \frac{1}{(s+1)!} \frac{1}{(t+1)!} \right] \quad s, t \geq 1 \quad (9)$$

Proof: See Appendix

1	2	3	4	5
0.625	0.275	0.079	0.017	0.003

Table 1: Rounded values of the probabilities of runs of lengths s from 1 to 5 according to eq.(5) .

Proposition 4 The correlation of two consecutive runs lengths is given by

$$\mathbb{E}(H_1 H_2 | E) - \mathbb{E}(H_1 | E) \mathbb{E}(H_2 | E) = 3 \left(\frac{1}{2} e^2 - 2e + \frac{5}{2} \right) - \left(\frac{3}{2} \right)^2 \approx 0.024 \quad (10)$$

Proof In order to compute the expectation $\mathbb{E}(H_1 H_2 | E)$ we use the formula

$$\mathbb{E}(H_1 H_2 | E) = \sum_{s_0, t_0=1}^{\infty} P(H_1 \geq s_0, H_2 \geq t_0 | E)$$

Using the telescopic form of the series we get

$$\begin{aligned} \sum_{t=t_0}^{\infty} P(H_1 = s, H_2 = t | E) &= 3 \left[\frac{1}{s+t_0+1} \frac{1}{s!} \frac{1}{t_0!} - \frac{1}{s+t_0+2} \frac{1}{(s+1)!} \frac{1}{t_0!} \right] \\ \sum_{t=t_0, s=s_0}^{\infty} P(H_1 = s, H_2 = t | E) &= 3 \left[\frac{1}{s_0+t_0+1} \frac{1}{s_0!} \frac{1}{t_0!} \right] \end{aligned}$$

To compute the last series we use that

$$\frac{1}{s_0+t_0+1} = \int_0^1 z^{s_0+t_0} dz$$

and so

$$\sum_{s_0, t_0=1}^{\infty} \frac{1}{s_0+t_0+1} \frac{1}{s_0!} \frac{1}{t_0!} = \int_0^1 dz \sum_{s_0, t_0=1}^{\infty} \frac{z^{s_0}}{s_0!} \frac{z^{t_0}}{t_0!} = \int_0^1 dz (e^z - 1)^2 = \frac{1}{2} e^2 - 2e + \frac{5}{2}$$

□

Using the above results the coefficient of correlation between the two variables is given by

$$\frac{\mathbb{E}(H_1 H_2 | E) - \mathbb{E}(H_1 | E) \mathbb{E}(H_2 | E)}{\sqrt{\mathbb{V}ar(H_1 | E) \mathbb{V}ar(H_2 | E)}} = \frac{2e^2 - 8e + 7}{8e - 21} \approx 0.0427 \quad (11)$$

3. An entropy-based runs test

Let X_1, \dots, X_n be a data series that is supposed to be i.i.d. and let H_1, \dots, H_N be the runs lengths extracted from it, where N denotes the number of runs.

The standard runs test commonly in use is based on a normal approximation of the number of runs N for n large:

$$\frac{N - (2n - 1)/3}{\sqrt{(16n - 29)/90}} \sim Z \quad (12)$$

where Z is the standard normal. Hence a level 0.05 test is usually performed computing the above Z-score and comparing it to the quantiles $z_{0.025} = -1.96$ and $z_{0.975} = 1.96$ [Sachs, 1982]. According to the result of the previous section, the lengths H_i are with good approximation a random sample extracted from the stationary distribution given by eq.(5). These probabilities for lengths from 1 to 5 are in the table 1. Let

$$g(s) = P(H_1 = s \mid E) / \sum_{s=1}^5 P(H_1 = s \mid E), \quad s = 1, \dots, 5$$

denote the distribution in eq. (5) truncated up to length 5.

Let $f_n(s)$ denote the empirical frequency of runs lengths up to length 5

$$f_n(s) = \frac{N_s}{N_1 + \dots + N_5} \quad s = 1, \dots, 5 \quad (13)$$

where N_s is the number of runs of length s in the data sequence of n elements. The test we consider here is based on the comparison of the empirical and the theoretical distributions f_n and g . The comparison is performed using the relative entropy or Kullback-Leibler divergence between two distributions (see for example Georgii [2008]).

In our setting the relative entropy of f_n with respect to g is given by

$$h_n = \sum_{s=1}^5 f_n(s) \log(f_n(s)/g(s)) \quad (14)$$

where $f_n(s) \log(f_n(s)/g(s)) = 0$ if $f_n(s) = 0$ and the natural logarithm is used. One has $h_n \geq 0$ and $h_n = 0$ if and only if $f_n(s) = g(s)$, $s = 1, \dots, 5$. Denoting $q_{0.95}$ the 0.95-quantile of the distribution of h_n , the test rejects the independence hypothesis if the value of h_n computed on a data sequence exceeds $q_{0.95}$. Tests based on relative entropy with rejection region of type $\{h_n > c\}$ have bounds of type I and type II errors that are exponentially small in the number of observations (Hoeffding's entropy test [Georgii, 2008]).

The quantiles of the distribution of h_n for a sequence of $n = 200, 500, 1000$ i.i.d. variables are obtained simulating 10,000 replicas using a pseudo random

n	1%	2.5%	5%	50%	95%	97.5%	99%
200	0.003355	0.003838	0.004390	0.013067	0.035700	0.041821	0.050466
500	0.000355	0.000721	0.001145	0.005619	0.014674	0.017268	0.020958
1000	0.000204	0.000390	0.000575	0.002707	0.007470	0.008468	0.009851

Table 2: Rounded values of the quantiles of the empirical distribution of h_n from simulated i.i.d. sequences of $n = 200, 500, 1000$ elements.

number generator (table 2). In this and in the following computations we have used the [R Development Core Team, 2008] free statistical software. We shall compare the powers of Z-score and entropy tests in the next section.

An alternative goodness of fit test could be the standard chi-square test, where the runs of length s observed are $Nf_n(s)$ and the expected are $Ng(s)$ for $s = 1, \dots, 5$. This approach has two drawbacks: the number of runs N itself is random and the usual condition on the number of observations per cell, $Nf_n(s) > 5$, requires a large n . Actually the runs of length 5 are expected to occur only 3 times in a sequence with $N = 1000$ runs (table1), corresponding to $n \sim 1500$, according to eq. (12).

Another runs test is based on the asymptotic normal approximation of the vector (N_1, \dots, N_s) . This test has been used for pseudo random number generators since it requires long series with $n \sim 4000$ [Knuth, 1988].

4. Power study

We investigate the performance of the entropy runs test in detecting dependence using as an alternative hypothesis models based on linear AR(1) processes:

$$X_t = \varphi X_{t-1} + Z_t$$

where Z_t is an i.i.d. sequence of standard normals. This model has a stationary univariate distribution for $\varphi \in (-1, 1)$ which is normal of mean 0 and variance $1/(1 - \varphi^2)$. For the properties of these models we refer to Brockwell and Davis [2002]. In our simulation initial datum is put to zero, and Z_t is simulated using a pseudo random number generator.

We have computed the power of the entropy runs test of level 0.05 for size $n = 500$ in function of the parameter φ in the interval $(-1, 1)$. For each value of the parameter the sequence has been replicated 500 times and the rejection rate was computed. Using the same sequences the power of the Z-score test has been also computed. The two powers are compared in fig. 1. Other simulations of AR(1) sequences with non normal (Cauchy) innovations show similar results.

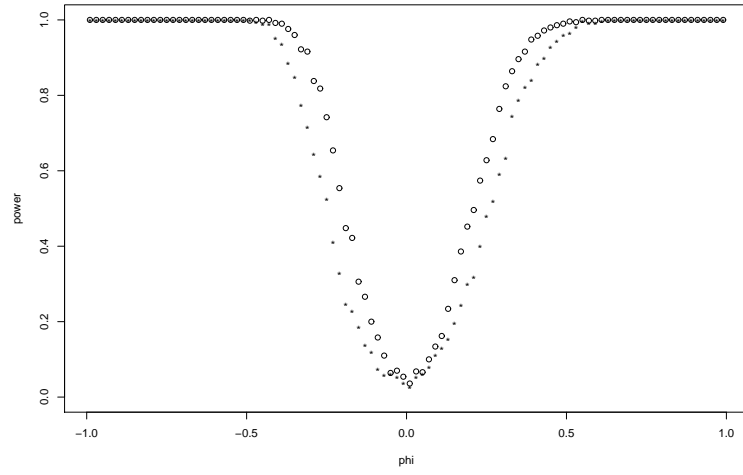


Figure 1: Powers of the runs tests of 0.05 level: Z-score (circles) and entropy (stars) as a function of the parameter φ of the AR(1) process with $n = 500$ elements.

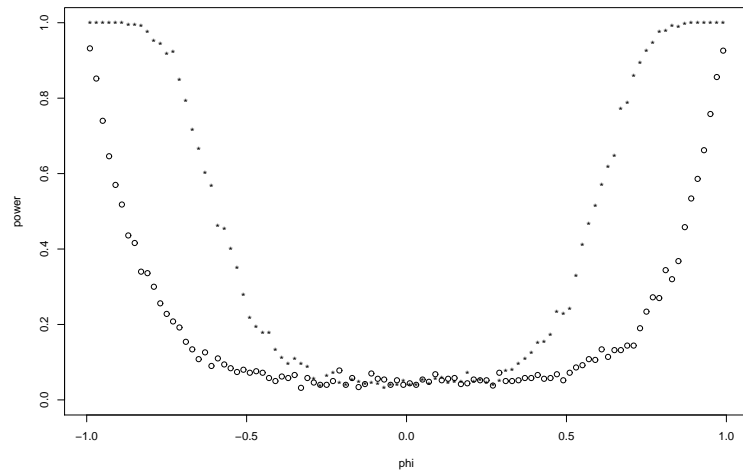


Figure 2: Powers of the runs tests of 0.05 level: Z-score (circles) and entropy (stars) as a function of the parameter φ of the non stationary process with $n = 500$ elements.

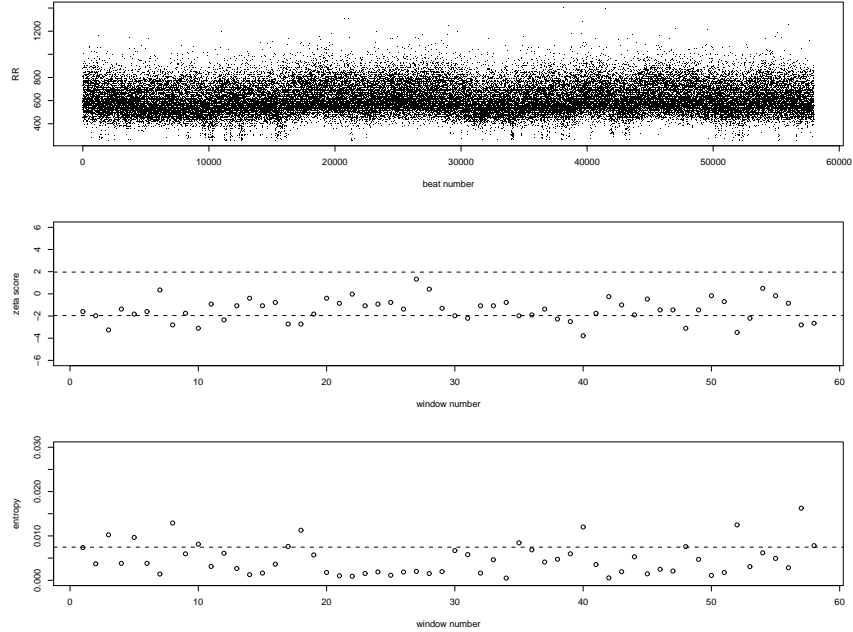


Figure 3: First panel: The time series of the RR intervals in msec. extracted from the 24 hours Holter ECG during atrial fibrillation. Second panel: The z-scores of the standard runs tests for windows of size 1000 and the acceptance region of level 0.05. Third panel: the runs entropy for the same windows and acceptance region of level 0.05.

We consider a second process obtained concatenating four AR(1) processes of size $n=125$: the first and the third one with φ , the second and the fourth with $-\varphi$. This process is not stationary, but its one dimensional marginal is stationary with respect to mean and to variance. **For each $\varphi \in (-1, 1)$ the process has been replicated 500 times and the powers of the two tests are computed. The powers are shown in fig. 2.**

For the stationary AR(1) process the power of entropy is slightly smaller than the other, in the non stationary case the former largely dominates the latter.

5. Application to atrial fibrillation

In the electrocardiogram the main peak (corresponding to systole) is labeled as R peak and the time intervals between two adjacent R peaks are called RR intervals. These intervals are inversely related to the instantaneous heart rate. The RR series shows a strong variability, which is known as heart rate variability (HRV)[Goldberger et al., 2000]. This variability is caused by several

factors the main of which is the neuroautonomic control, which acts accelerating and decelerating the heart rate. Hence the analysis of the runs up and down of this series may be a tool for investigating the efficiency of this control.

Atrial fibrillation is a commonly diagnosed arrhythmia which affects up to 1 % of the general population. In this pathology the conduction of the electrical stimulation through the atria is irregular leading to a corresponding sequence of irregular RR intervals, which are no longer controlled by the neuroautonomic system.

In subjects with atrial fibrillation the RR series shows very large uncorrelated fluctuations if compared to normals and other pathologies (see for instance Urbanowicz et al. [2007] and references therein).

Although the irregular rhythm during atrial fibrillation is generally accepted to be random, rigorous statistical testing are still lacking.

In Cammarota and Rogora [2005] data analysis performed using parametric tests on a small group of subjects showed that the hypothesis of serial independence is not rejected in the large majority of time; in the remaining the series can be modeled as an auto regressive model (AR) of very short memory.

This type of data have two features that make useful non parametric tests: the distribution is not normal, having a longer right tail; there are outliers due to the automated detection of the R peaks. Furthermore adjacent data values are generally largely different so the frequency of ties is negligible and the runs are well defined. We apply both standard and entropy runs tests in one case of atrial fibrillation.

This series was exported by the native software of a Rozinn Holter (Glen-dale, USA) equipment for ambulatory diagnostics in Policlinico of La Sapienza University of Rome.

This case shows an episode of paroxysmal atrial fibrillation, a condition in which the irregular rhythm occurs and terminates in a non-predictable way over time having duration from a few minutes up to days. In the case under analysis the episode of fibrillation was consisting of about 58,000 beats; extraction of these beats produces an RR time series of the same length (fig. 3, first panel). We have divided the series into 58 consecutive disjoint windows of size $n = 1000$. The entropy runs test is applied to each window and the values of the statistic are reported in fig. 3 (third panel) and compared to the Z-score of the standard runs test (second panel).

The same procedure was repeated for windows of size $n = 500$ and $n = 200$ and the results are reported in table 3. The numbers of widows in which the hypothesis is not rejected is near the same for sizes $n = 200, n = 500$. In case $n = 1000$ the acceptance rates (38/58 and 46/58) are smaller and their difference is larger. Both these facts can be explained since longer data windows could be non stationary with respect to mean. Finally since the third column is less than the fourth, the entropy seems to be a more conservative test, which is in agreement with fig. 1.

n	windows	Z-score	entropy
200	290	258	260
500	116	97	102
1000	58	38	46

Table 3: Acceptance rates of the Z-score test (third column) and of the entropy test (fourth column) for 3 sizes windows in the time series of atrial fibrillation. First column: size of the window; second: number of windows.

6. Conclusions

The first result states that the correlation of two consecutive runs lengths is small. This allows to use all the runs extracted from a data series as an approximately random sample, avoiding to eliminate runs or data values from the series. The second result concerns an entropy-based runs test that uses as statistic the relative entropy of the empirical distribution of the runs lengths $f_n(s) = N_s/N$ with respect to the theoretical one in the i.i.d. hypothesis. This test is compared to the Z-score runs test, based on the asymptotic normal approximation of the total number of runs N .

The powers of the two tests have been compared using as alternative processes based on AR(1) models with parameter $\varphi \in (-1, 1)$ which describes the strength of the interaction. The results are different if the process is stationary or not. In the stationary case the power of the entropy is smaller than the other, while in the non stationary the opposite happens. Power studies were lacking. This is a very preliminary investigation and extensive studies are needed to get some conclusion. We remark that the Z-score depends only on the number N of runs, while the entropy is less sensitive on N , since the empirical frequencies are computed just normalizing with respect to this number. For this reason a compound test which uses both the informations should be of interest.

The entropy runs test has been applied to time series of inter beat duration in atrial fibrillation. These series are conjectured to be independent basing on physiologic arguments and from previous investigations mainly without tests of statistical significance. The runs up and down of the RR series can be interpreted as decelerations and accelerations of the heart rate, hence their departure from the i.i.d. can reflect the control of the neuroautonomic system.

Both the Z test and the entropy do not reject the independence hypothesis for a large majority of the time. The rate of acceptance of entropy seems to be larger; both the rates decrease as the size of the

window increases, but the entropy seems more stable. These facts depend on several factors; one of them could be the slowly varying mean typical of physiologic signals. In the windows in which the independence is rejected an AR model of short memory was conjectured [Cammarota and Rogora, 2005], and the problem deserves a detailed analysis. The fraction of the windows in which the independence is not rejected could be used in clinics as an index of the strength of the pathology.

Some other questions are naturally to be considered. The entropy test is based on numerical evaluation of the distribution of the statistic and relies upon the capability of pseudo random generators to simulate i.i.d. sequences. Investigations on the asymptotic distribution of the statistic can take advantage of the small correlation between two consecutive runs, a result that has not been fully exploited here. Finally both tests assume that the variables are continuous, but in data series this is only an approximation mainly for low measurement resolution.

7. Appendix

Proof of Proposition 2. The proof of eq. (5) goes back to Levene and Wolfowitz [1944]; a recent exposition is in Cammarota and Curione [2008]; a proof obtained from the bivariate distribution is given below. Eq. (6), a rewriting of eq. (2), can be obtained from the distribution (5) as follows:

$$\mathbb{E}(H_1 | E) = 3 \sum_{s_0=1}^{\infty} P(H_1 \geq s_0 | E)$$

$$P(H_1 \geq s_0 | E) = \sum_{s=s_0}^{\infty} \frac{1}{s+2} \frac{1}{s!} - \frac{1}{s+3} \frac{1}{(s+1)!} = \frac{1}{s_0+2} \frac{1}{s_0!}$$

where the last equality is obtained summing the telescopic series. Using that

$$\frac{1}{s_0+2} = \int_0^1 dz z^{s_0+1}$$

one gets eq. (6) since

$$\sum_{s_0=1}^{\infty} \frac{1}{s_0+2} \frac{1}{s_0!} = \int_0^1 dz z(e^z - 1) = \frac{1}{2}$$

Let us now compute the second moment $\mathbb{E}(H_1^2 | E)$. This is given by

$$3 \sum_{s=1}^{\infty} s^2 \left[\frac{1}{s+2} \frac{1}{s!} - \frac{1}{s+3} \frac{1}{(s+1)!} \right]$$

Using $s^2 = (s+1)^2 - 2s - 1$ in the second contribution we get

$$\frac{s^2}{(s+2)s!} - \frac{s^2}{(s+3)(s+1)!} = \frac{s^2}{(s+2)s!} - \frac{(s+1)^2}{(s+3)(s+1)!} + \frac{2s+1}{(s+3)(s+1)!}$$

The first two terms form a telescopic series whose sum is $1/3$. The third one gives

$$\frac{2(s+1)}{(s+3)(s+1)!} - \frac{1}{(s+3)(s+1)!} = \frac{2}{(s+3)s!} - \frac{1}{(s+3)(s+1)!}$$

which can be treated as before and gives after summation $2e - 29/6$. \square

Proof of Proposition 3. A complete proof is contained in the paper Levene and Wolfowitz [1944]. Here we explain the idea of the proof in the particular case $H_1 = 3, H_2 = 3$ to avoid notational complications. Conditioning to E is the same as to B and this gives the factor 3. Conditioning to B means that 2 is a minimum, 5 is a maximum and 8 is a minimum. One has, denoting f the density of the X_i ,

$$\begin{aligned} P(H_1 = 3, H_2 = 3 | E) &= \\ &= 3 P(X_1 > X_2, X_2 < X_3 < X_4 < X_5, X_5 > X_6 > X_7 > X_8, X_8 < X_9) \\ &= 3 \int_{-\infty}^{+\infty} dx_5 f(x_5) P(X_1 > X_2, X_2 < X_3 < X_4 < x_5, x_5 > X_6 > X_7 > X_8, X_8 < X_9) \\ &= 3 \int_{-\infty}^{+\infty} dx_5 f(x_5) P(X_1 > X_2, X_2 < X_3 < X_4 < x_5) P(x_5 > X_6 > X_7 > X_8, X_8 < X_9) \end{aligned}$$

where the last equation follows by independence. One also has

$$\begin{aligned} P(X_1 > X_2, X_2 < X_3 < X_4 < x_5) &= P(X_2 < X_3 < X_4 < x_5) - P(X_1 < X_2 < X_3 < X_4 < x_5) \\ P(x_5 > X_6 > X_7 > X_8, X_8 < X_9) &= P(x_5 > X_6 > X_7 > X_8) - P(x_5 > X_6 > X_7 > X_8 > X_9) \end{aligned}$$

Denoting F the distribution function of the X_i the first contribution is

$$P(X_2 < X_3 < X_4 < x_5) = \int_{-\infty}^{x_5} dx_4 f(x_4) \int_{-\infty}^{x_4} dx_3 f(x_3) \int_{-\infty}^{x_3} dx_2 f(x_2) = \frac{1}{3!} F^3(x_5)$$

Using similar expressions for the other contributions we get

$$\begin{aligned} P(H_1 = 3, H_2 = 3 | E) &= 3 \int_{-\infty}^{+\infty} dx_5 f(x_5) \left[\frac{1}{3!} F^3(x_5) \frac{1}{3!} F^3(x_5) - \frac{1}{3!} F^3(x_5) \frac{1}{4!} F^4(x_5) \right. \\ &\quad \left. - \frac{1}{4!} F^4(x_5) \frac{1}{3!} F^3(x_5) + \frac{1}{4!} F^4(x_5) \frac{1}{4!} F^4(x_5) \right] \end{aligned}$$

The integration of the first contribution gives

$$\int_{-\infty}^{+\infty} dx_5 f(x_5) \frac{1}{3!} F^3(x_5) \frac{1}{3!} F^3(x_5) = \frac{1}{7} \frac{1}{3!} \frac{1}{3!}$$

Since the other contributions are similar, this completes the proof of the particular case. \square

Proof of eq.(5) We compute the first marginal of eq.(9). The first two contributions in eq.(9) form a telescopic series with respect to t :

$$\sum_{t=1}^{\infty} \frac{1}{s+t+1} \frac{1}{s!} \frac{1}{t!} - \frac{1}{s+t+2} \frac{1}{s!} \frac{1}{(t+1)!} = \frac{1}{s+2} \frac{1}{s!}$$

The same for the third and fourth:

$$\sum_{t=1}^{\infty} \frac{1}{s+t+2} \frac{1}{(s+1)!} \frac{1}{t!} - \frac{1}{s+t+3} \frac{1}{(s+1)!} \frac{1}{(t+1)!} = \frac{1}{s+3} \frac{1}{(s+1)!}$$

Hence

$$\sum_{t=1}^{\infty} P(H_1 = s, H_2 = t | E) = 3 \left[\frac{1}{s+2} \frac{1}{s!} - \frac{1}{s+3} \frac{1}{(s+1)!} \right]$$

\square

Acknowledgments.

We thank Mario Curione (Clinical Science Department, University of Rome ‘La Sapienza’) for providing data and for valuable discussions. **We thank the referees for stimulating criticism.**

References

- N. Balakrishnan and M.V. Koutras. *Runs and Scans with Applications*. Wiley, 2001.
- J.V. Bradley. *Distribution-free statistical tests*. Prentice-Hall, 1968.
- L. Breiman. *Probability*. SIAM, 1993.
- P.J. Brockwell and R.A. Davis. *Introduction to time series and forecasting*. Springer-Verlag, 2002.
- C. Cammarota and M. Curione. Analysis of extrema of heartbeat time series in exercise test. *Mathematical medicine and biology*, 25:87–97, 2008. ISSN 1477-8599. doi: 10.1093/imammb/dqn005. URL <http://dx.medra.org/10.1093/imammb/dqn005>.
- C. Cammarota and E. Rogora. Independence and symbolic independence of nonstationary heartbeat series during atrial fibrillation. *Physica. A*, 353:323–335, 2005. ISSN 0378-4371.
- E.J. Chen and W.D. Kelton. Determining simulation run length with the run test. *Simulation Modelling Practice and Theory*, 11:237–250, 2003.

- H-O. Georgii. *Stochastics*. de Gruyter, 2008.
- A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000. Circulation Electronic Pages: <http://circ.ahajournals.org/cgi/content/full/101/23/e215>.
- D.E. Knuth. *The art of computer programming, Vol. 2*. Addison-Wesley, 1988.
- H. Levene and J. Wolfowitz. The covariance matrix of runs up and down. *Annals Math. Stat.*, 15:58–69, 1944.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- V. K. Rohatgi. *An introduction to Probability Theory and Mathematical Statistics*. Wiley, 1976.
- L. Sachs. *Applied Statistics*. Springer, 1982.
- K. Urbanowicz, J.J. Zebrowski, R. Baranowski, and J.A. Holyst. How random is your heart beat? *Physica A*, 384:439–447, 2007.
- J. Wolfowitz. Asymptotic distribution of runs up and down. *Annals Math. Stat.*, 15(2):163–172, 1944.