
Joseph Horowitz, Ph.D.
Department of Mathematics and Statistics
Control in Biomedical Systems Research Laboratory
University of Massachusetts
Amherst, MA 01003

Anand (and others) – Here are some comments on your paper, which is referred to as [D] below. The goal stated in the Conclusions section of the paper, to “provide a potentially powerful universal quantitative framework to explain the formation and destruction of rotational events in mammalian cardiac fibrillation” is a worthy goal, but in my opinion only a first step toward it has been achieved in [D]. I hope you will understand that the criticisms below are intended to be friendly and constructive. In addition to the comments, a suggestion for a more comprehensive model is included.

Notation $F_1, F_2, \ldots$ will denote the interformation times of the 1st, 2nd, etc., phase singularities (PSs), thus $T_1 = F_1, T_2 = F_1 + F_2, T_3 = F_1 + F_2 + F_3, \ldots$ are the actual “birth times” (times of formation) of the successive PSs, and $D_1, D_2, D_3, \ldots$ are the “lifetimes” (time to destruction) of the successive PSs.

A. Poisson process is not demonstrated The paper [D] asserts that $F_n$ is the sequence of interarrival (or, in the present context, interformation) times of a Poisson process, and similarly for the sequence $D_n$. Using histograms, nonlinear regression, $\chi^2$, etc., it is shown that both the $F_n$ and $D_n$ sequences follow exponential distributions, with parameters $\lambda_d$ and $\lambda_b$, respectively. This is consistent with underlying Poisson processes, but is insufficient to establish “Poissonness”. (Note: The word ‘renewal’ is generally omitted, so “Poisson process” means the Poisson renewal process, cf. Resnick, “Adventures”, p. 211.) It is also necessary to show that the random variables (rvs) $F_n$ are independent, identically distributed (iid) (cf. “Adventures”, p. 7); and likewise for the sequence $D_n$. The histograms, etc., in [D] that demonstrate the exponential distributions are OK as far as they go, but they do not prove independence. In fact, it is possible for all the rvs $F_n$ to have the same exponential ($\lambda_d$) distribution, but to be dependent. (A similar comment pertains to $D_n$.) Example: let $X_n$ be an iid sequence of rvs with the Gamma(1/2, 1) distribution, i.e., gamma distribution with shape and scale parameters $1/2$ and $1$. Let $Y_n = \frac{1}{2}(X_n + X_{n+1})$ be an order 2 moving average of the $X_n$ sequence. Since the $X_n$ are independent gammas with the same scale, $X_n + X_{n+1}$ is gamma(1,1), and $Y_n = \frac{1}{2}(X_n + X_{n+1})$ is gamma(1,1/2), which is the same as exponential ($\lambda = 2$). Thus $Y_n$ is by construction a sequence of identically distributed, but dependent exponential rvs, which therefore does not arise as the interarrival times of a Poisson process. If you apply histograms, nonlinear least squares fitting, etc., the sequence $Y_n$ will “pass” these tests – they will show an exponential distribution, but a check of the autocorrelation function will exhibit nonzero lag 1 correlation. I have done this (in R 4.0.0) and can send the results. The example is easily modified for any value of $\lambda$. There are many ways to check independence, both graphical and by formal statistical tests, in the statistical and time series literature. If the rvs $F_n$ are not independent, then we are not in the realm of renewal theory; moreover, as pointed out below, renewal theory is not relevant to $D_n$ in any case. If the $F_n$ are not independent, one can ask whether they form a stationary sequence. In any case, what does independence or lack thereof mean in terms of electrophysiology? (Same questions for $D_n$.)

B. Asymptotic slope The asymptotic slope property of $N(t)$, the number of renewals during (0,t], illustrated in Fig. 1, holds not only for the Poisson process, but for all renewal processes (“Adventures”,
p. 189), for all stationary point processes (Daley and Vere-Jones, An Introduction to the Theory of Point Processes, vol. I, p. 60), and some others.

C. No physical meaning of $D_1 + \ldots + D_n$ The formation times $T_n = F_1 + \ldots + F_n$ have a meaningful linear order, whereas the sums $D_1 + \ldots + D_n$ have no physical meaning, because the lifespans of two or more PSs may overlap. Thus renewal theory is not relevant to the $D_n$. It is reasonable to ask whether the $D_n$ are iid, and it is shown in [D] that they are “id” with exponential ($\lambda_n$) distributions, but independence must be checked, as noted in A.

D. MaxEnt does not establish exponential distribution for these data The result in S13 (proven there only for discrete distributions) is a general result that does not tie the exponential distribution to the data in any way. The same argument could be made for any set of data, whether from an exponential distribution or not, assuming the restriction on the expected value. (A proof for continuous distributions is in Rao, Linear Statistical Inference and its Applications, 2nd ed., p. 173.) References 35 and 42 of [D] show that, if you assume certain constraints, MaxEnt will give particular distributions, but it is necessary to show that the assumed constraints are required or at least plausible for the data at hand; that has not been done. Thus the last sentences in the sections Comparison With Maximum Entropy Predicted Distribution and Origins of the Exponential Limiting Distribution—Applying the Principle of MaxEnt are called into question.

E. Suggested model As laid out in [D], the sequences of rvs, $F_n$ and $D_n$, are hardly related to each other or to the underlying physiology; in particular no connection is drawn to the spatial distribution of the PSs. Further, it is not clear how the issues raised in Future Directions could be addressed via these two sequences. I suggest below a more comprehensive model that unifies everything in one package and which allows a broad range of questions to be posed and (maybe!) answered in a principled, rigorous manner. Two slightly different versions of the model are proposed; at this point I’m not sure which, if either, is better or if each can be used for different questions. Recall the PS formation times $T_n$ above, so $F_n = T_n - T_{n-1}$ (defining $T_0 = 0$). The technical tool is marked point processes (treated briefly in “Adventures” section 4.10, more thoroughly in Daley and Vere-Jones, op. cit.). These models simply organize the data in a unified way.

Model I The set of times $T_1 < T_2 < \ldots$ can be considered as a point process (the “ground process”) on the nonnegative real line $[0, \infty)$. To each point $T_n$ we associate a 3-dimensional “mark” consisting of the location $(X_n,Y_n)$ and the lifetime $D_n$ of the PS that is born at time $T_n$. Thus the mark is the (random) point $(X_n,Y_n,D_n)$ in 3D. Here the locations are taken as 2D, but we could as well have a 3D location $(X_n,Y_n,Z_n)$, hence a 4D mark $(X_n,Y_n,Z_n,D_n)$. The locations could even be on a more complicated geometric object such as a manifold, depending on the model of the heart. As far as I could tell, the locations are available, but I may have misunderstood this point.

Model IA If the locations are not available, the ground point process would again be $T_n$, but the marks would simply be $D_n$. This would at least treat $F_n$ and $D_n$ in a unified way.

Model II The ground process in this case is the spatiotemporal point process $(X_n,Y_n,T_n)$, and the corresponding mark is $D_n$. Again, the locations could be 3D, making the ground process 4D, but the marks remain 1D.

F. Some questions that can be approached within these models
1) Are the ground process and the marks independent? This is the same as asking whether the spatial locations and lifespans of the PSs are independent of the formation times. In model IA the question becomes whether the formation times and the lifespans are independent.

2) Are the spatial locations and the marks independent?

3) Are the spatial locations and formation times independent?

In each case, if independence does not hold, what is the nature of the relationship between the two entities and, either way, what is the physiological significance? My own intuition is that independence will not hold in these cases because everything is strongly electrically connected; on the other hand, in the context of fibrillation, perhaps any such dependence would be “washed out”. Another question is what is the electrophysiological import of the parameters of the distributions of $F_n$, $D_n$, etc. These seem to me to be interesting questions for electrophysiology, but I am not qualified to make such a judgment.

Some of these questions can be better formulated in model I or IA, some in model II. There is a substantial literature on marked point processes, and some of the above problems can be addressed with off-the-shelf methods, whereas others will require development of new approaches.

G. Odds and ends

1. The proof in S1 is incorrect: the sum labeled “pmf of geometric distribution” is not the pmf, it is $= 1$. Similarly for the integral labeled “pmf of the exponential distribution”. The proof can be easily repaired.

2. $\lambda$ is labeled $\lambda\%$ throughout. I didn’t understand the meaning of % here.

3. The values of $\lambda$ and of the means in Table 3 should be $> 0$.

4. The function in eq. (2), usually denoted by $F(t)$, is the cumulative distribution function, and is the probability that formation will occur before time $t$.

5. The term ‘cross-validation’ (in the section on sensitivity analysis) has a specific technical meaning in statistics, not as used here.

6. The section Implications for Overall Cardiac Fibrillation System Dynamics raises some interesting questions, especially that of regeneration and termination. There are so-called regenerative processes (“Adventures”, section 3.12) and terminating renewal processes (“Adventures”, ch. 3) that may be relevant.