

A parsimonious alternative to the pacemaker/accumulator process in animal timing

Alliston K. Reid ^{a,*}, David L. Allen ^b

^a *Department of Psychology, Wofford College, 429 North Church Street, Spartanburg, SC 29303-3663, USA*

^b *Department of Mathematics, Eastern Oregon University, La Grande, Oregon, USA*

Received 7 January 1998; received in revised form 1 July 1998; accepted 4 July 1998

Abstract

Models of animal timing often assume the existence of a pacemaker whose variance is produced by a Poisson process. Later stages in these models typically include an accumulator, additional sources of variance, and a process in which the accumulated value is compared to reference memory. Together, these stages produce a constant coefficient of variation (i.e. the ‘scalar property,’ often observed in experiments) even though at least one of the sources of variance is not scalar. We propose a more parsimonious source of scalar variance. We define a linear failure rate function and demonstrate that its resulting distribution function, called the Rayleigh distribution, always produces the scalar property without requiring: (a) additional sources of variance, or (b) that comparisons with reference memory always be based on ratios. Continuous and discrete versions of this model are derived. This model may serve as a more parsimonious pacemaker than others currently proposed and may replace the pacemaker/accumulator process in timing models. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Animal timing; Cognitive processes; Pacemakers; Stochastic models

1. Introduction

Most pacemaker/accumulator models of animal timing assume a Poisson process, which produces a constant mean rate of pulses and exponentially distributed interarrival times. These pacemakers are parsimonious because they are memoryless. In

the information-processing version of scalar expectancy theory (SET) (Gibbon, 1977; Gibbon and Church, 1984) these pulses pass through a switch into an accumulator. The switch produces an additional source of variance at the beginning and end of the switch closure. The accumulation of these pulses produces a value that increases linearly with time. Finally, this linearly increasing value is compared to reference memory. The comparison process operates on ratios which produces

* Corresponding author. Tel.: +1 864 5974642; fax: +1 864 5974640; e-mail: reidak@wofford.edu

the multiplicative variance called the ‘scalar property.’ Therefore, “[it] means in essence that a source of random variability, not in itself scalar, induces scalar expression of that variability at different target times because the mechanism underlying performance entails a rescaling of the random variable by the target time” (Gibbon, 1991, p. 20). SET has additional sources of the scalar property. For example, the memory source in which reinforced accumulator values are encoded and subsequently decoded involves multiplicative variance, as does the threshold for a ratio decision mechanism.

We argue that an alternative pacemaker may be more parsimonious than this pacemaker/switch/accumulator system. We demonstrate below that a simple pacemaker, whose variance is based solely on a linear failure rate function, produces the equivalent of the linear relation with time, as well as a constant coefficient of variation at all interval values. That is, this simple process produces the scalar property even without subsequent ratio comparisons to reference memory, a threshold for the decision mechanism, or a scalar encode/decode mechanism. In the sections that follow, we derive the properties of continuous and discrete versions of this pacemaker.

2. Sample problem

When developing probabilistic models of behavior, a specific example often helps make the problem more intuitive and understandable. We developed this model while analyzing data collected from a fixed consecutive number (FCN) schedule of reinforcement, and it may be a helpful example. Using pigeons as subjects, an FCN schedule (Mechner, 1958) is a discrete-trials procedure that requires the pigeon to peck the left key N times or more, before terminating the trial by a single peck on the right key. Food is delivered if the number of left pecks is greater than or equal to the required ratio value (N). The counter is reset for the next trial. Trials ending with less than N left pecks reset the counter and typically produce timeout. The effects of the timeout will not be considered in this analysis.

The general problem to be solved is predicting how many left pecks will occur in any trial. Because pigeons may not be able to count accurately, they may solve this problem by pecking at a fairly constant rate, and use their sense of the passage of time to determine when to peck the right key. Therefore, this example provides us with the opportunity to consider both discrete (peck-based) implementations and continuous (time-based) implementations of our probabilistic model. We begin with the continuous case.

3. The continuous case

The continuous case is naturally viewed in the context of survival analysis, a topic within applied probability theory. In psychology, survival analysis has been extensively applied to the analysis of response times (Townsend and Ashby, 1983; Luce, 1986; Balakrishnan, 1994). In survival analysis, one studies the ‘lifetime’ of ‘individuals’ where these terms are interpreted broadly. The individual could be a component in a system with its lifetime being the time until it fails. In our situation, the individual is a pigeon trial, and the lifetime is the duration of the trial (the time until the pigeon pecks the right key). In the discrete model (considered later), the time would be measured in key pecks, and in the continuous model it would be the time since the beginning of the trial. Ross (1980) and Ross (1984) are good introductory sources for background material on probability theory and survival analysis. Cox and Oakes (1984) provide a thorough discussion of survival analysis.

4. Failure rate functions

Probability distributions in the context of survival analysis are usually described in terms of the failure (or hazard) rate function, $\lambda(t)$. Let Z be a positive continuous random variable, intended to represent the duration of a trial (equivalently, the lifetime of an individual). Let $F(t)$ be the distribution function of Z , that is, $F(t) = P\{Z \leq t\}$ and let $f(t)$ be the density function of Z , that is

$f(t) = F'(t)$. The failure rate function $\lambda(t)$ for Z , also called the hazard rate function or the intensity function, is defined by:

$$\lambda(t) = \frac{F'(t)}{1 - F(t)} = \frac{f(t)}{1 - F(t)} \quad t > 0. \quad (1)$$

The failure rate function uniquely determines the distribution function and vice versa.

Intuitively, the failure rate function $\lambda(t)$ represents the conditional probability density that an individual will fail, given it has survived until time t . In our sample problem, this is equivalent to the conditional probability density that the pigeon will peck the right key, given that it has reached this point in the trial. Specifically, let Δt be a small positive number, then the probability of failure in the next Δt amount of time, given survival up until t , is:

$$\begin{aligned} P\{Z \in (t, t + \Delta t) | Z > t\} &= \frac{P\{Z \in (t, t + \Delta t)\}}{P\{Z > t\}} \\ &\approx \frac{f(t)\Delta t}{1 - F(t)} = \lambda(t)\Delta t, \end{aligned} \quad (2)$$

which follows directly from the definition and is derived in Ross (1980). If the failure rate function $\lambda(t)$ were constant, then Z would be exponentially distributed and represent the arrival rate in a Poisson process. The time until the first Poisson arrival is exponentially distributed and could be interpreted as the time until failure (switching) in the context of survival analysis. Most pacemaker/accumulator models of animal timing assume $\lambda(t)$ is constant, which produces a constant mean rate of pulses. A pacemaker with constant mean rate is the simplest, most parsimonious, stochastic pacemaker possible because it assumes every pulse is independent of all preceding pulses, producing a memoryless process.

Nevertheless, the accumulation of these pulses produces a value that increases linearly with time. For example, in the information-processing version of SET, pulses from the pacemaker pass through a switch (with its own source(s) of variability) into an accumulator. The accumulated value increases linearly with time and is compared to the value stored in the reference memory. The comparison process operates on ratios, which produces the multiplicative variance commonly called

the ‘scalar property.’ Therefore, SET predicts a constant coefficient of variation even though the variance of the pacemaker is not scalar but Poisson.

5. An alternative pacemaker

As an alternative way of modeling the time until the pigeon switches under an FCN schedule of reinforcement, it seems reasonable to assume the probability of switching (failure) is initially low and increases with time. A natural choice is the linear relation $\lambda(t) = mt$, where the parameter m satisfies $m > 0$. From this failure rate function, one may derive the distribution function as shown in Ross (1980). The resulting distribution function is:

$$F(t) = 1 - \exp\left[-\frac{1}{2}mt^2\right], \quad t \geq 0, \quad (3)$$

known as the Rayleigh distribution. The derivative of this distribution is its density function:

$$f(t) = mt \exp\left[-\frac{1}{2}mt^2\right], \quad t \geq 0. \quad (4)$$

Fig. 1 depicts three probability density functions with different means and their corresponding distribution functions.

Using the fact (obtained from tables of integrals) that $\int_0^\infty x^2 e^{-x^2} dx = \sqrt{\pi}/4$, we find:

$$\text{Expected value} = E[Z] = \sqrt{\frac{\pi}{2m}}, \quad (5)$$

and:

$$\text{Variance} = \text{Var}[Z] = \frac{4 - \pi}{2m}. \quad (6)$$

Combining these equations, the coefficient of variation is the constant:

$$\frac{\sigma}{\mu} = \sqrt{\frac{4 - \pi}{\pi}} \approx 0.523. \quad (7)$$

Since the coefficient of variation is the same for any choice of the parameter m , the Rayleigh distribution is said to have the scalar property.

The Rayleigh distribution has several interesting properties. The family of Rayleigh distributions is closed under scalar multiplication. That is, if Z has a Rayleigh distribution with parameter m ,

then, for any $c > 0$, cZ has a Rayleigh distribution with parameter mc^{-2} . Also, if $Y = Z^2$, then Y is exponentially distributed with parameter $m/2$, which can be seen as follows:

$$P\{Y > t\} = P\{Z^2 > t\} = P\{z > \sqrt{t}\} \\ = \exp(-\frac{1}{2}m\sqrt{t^2}) = \exp\left(-\frac{m}{2}t\right). \tag{8}$$

The survivor function, $\bar{F}_Z(t) = 1 - F_Z(t) = P\{Z > t\}$, for the Rayleigh decays exponentially with the square of time:

$$\bar{F}_Z(t) = 1 - F_Z(t) = \exp(-\frac{1}{2}mt^2). \tag{9}$$

Finally, the Rayleigh has a property related to the memoryless property of the exponential dis-

tribution. The memoryless property states that, conditional on survival up to k , the probability that survival continues for at least l units is the same as the unconditional probability of survival for at least l units. More formally, a random variable X with the memoryless property satisfies:

$$P\{X > k + l | X > k\} = P\{X > l\}.$$

The Rayleigh distribution has the following distributional property that has obvious similarities to the memoryless property:

$$P\{Z > k + l | Z > k\} = e^{-mkl} P\{Z > l\} \tag{10}$$

6. The discrete case

We will construct a discrete model analogous to the continuous model presented above. The construction is based on the Rayleigh distribution with $\lambda(t) = mt$ although one could use the same construction with any failure rate function. The discrete case is not as analytically tractable as the continuous case. However, it does ‘approximately’ share the same properties in the sense that, as a random variable, the discrete case approximates the continuous case. This statement will be apparent from the construction below. The construction involves two stages of approximation. First, we introduce Y as a continuous random variable that approximates a Rayleigh variable Z . Secondly, we approximate the continuous variable Y by a discrete variable R . The distribution of R , which we will call the discrete Rayleigh distribution, will closely resemble the continuous Rayleigh distribution.

Let Y be a random variable with failure rate function defined by $\lambda(t) = km$ for $t \in [k - 1, k)$ where $m > 0$ is constant and $k = 1, 2, \dots$. Observe $\lambda(t)$ is a step function, constant on each interval $[k - 1, k)$, and is approximately linear with slope m . Thus it is reasonable to expect Y to approximate a Rayleigh variable with parameter m . Now, define R to be Y rounded up to the nearest integer, that is, $R = \min\{n | n \in \text{Integers and } n > Y\}$. The distribution function for R , $F_R(k)$ is obtained as follows. Let k be any positive integer, then:

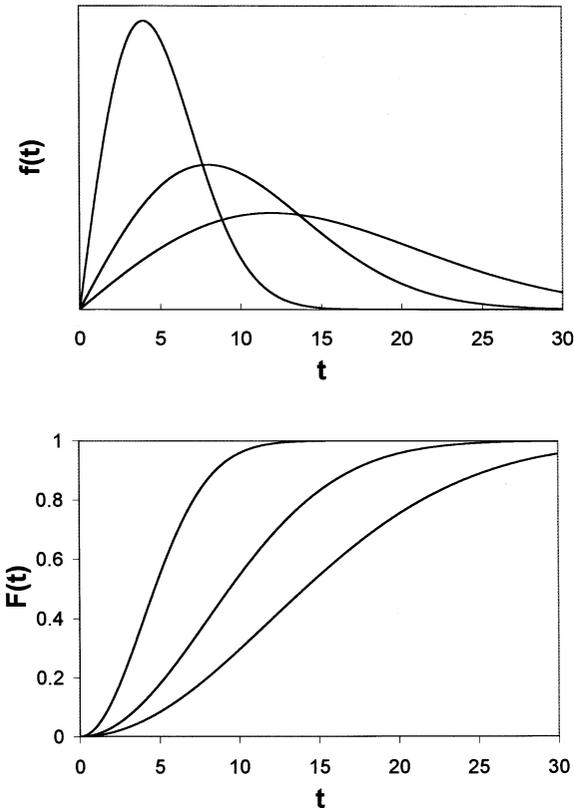


Fig. 1. Probability density functions, $f(t)$, (top panel) and cumulative distribution functions, $F(t)$, (bottom panel) for the Rayleigh family with means at 5, 10, and 15. The slopes of the linear failure rate functions were $m = 0.062832$, 0.015708 , and 0.006981 .

$$\begin{aligned}
 1 - F_R(k) &= P\{R > k\} = P\{Y > k\} \\
 &= 1 - F_Y(k) = \exp\left(-\int_0^k \lambda(t) dt\right) \\
 &= \exp\left(-\int_0^1 \lambda(t) dt - \int_1^2 \lambda(t) dt\right) - \\
 &\dots - \int_{k-1}^k \lambda(t) dt \\
 &= \exp(-m - 2m - \dots - km) \\
 &= e^{-m}(e^{-m})^2 \dots (e^{-m})^k \\
 &= \beta\beta^2 \dots \beta^k
 \end{aligned}$$

where we have introduced the notation $\beta = e^{-m}$. Next we submit an interpretation for this discrete random variable.

Consider a sequence of independent coin tosses (Bernoulli trials) where the probability of heads at each toss is allowed to vary. Let p_k be the probability of heads on the k th toss, and let $q_k = 1 - p_k$ be the probability of tails. Let M be the number of tosses until the first head. For M to equal k , there must be $k - 1$ tails followed by a head, thus $P\{M = k\} = q_1q_2 \dots q_{k-1}p_k$. For M to exceed k , the first k tosses must all be tails, thus $P\{M > k\} = q_1q_2 \dots q_k$. If we let $q_k = \beta^k$ then M has exactly the same distribution as the discrete Rayleigh, R , in the preceding paragraph. If $q_k = q$, constant for all k , the result is the memoryless geometric distribution, analogous to the continuous case where a constant failure rate produces the memoryless exponential distribution. The discrete Rayleigh has the same similarity to the memoryless property as the continuous Rayleigh. With e^{-m} replaced by β in Eq. (10):

$$P\{R > k + l | R > k\} = \beta^{kl}P\{R > l\}. \tag{11}$$

Fig. 2 compares the shapes of the discrete and continuous Rayleigh distributions. Table 1 compares the descriptive statistics for the continuous and corresponding discrete Rayleigh distributions with parameters chosen to match their means. It is evident that the discrete Rayleigh distribution approximates the Rayleigh distribution. The probability mass function for the discrete Rayleigh has the same shape as the density of the Rayleigh distribution. Unfortunately, the

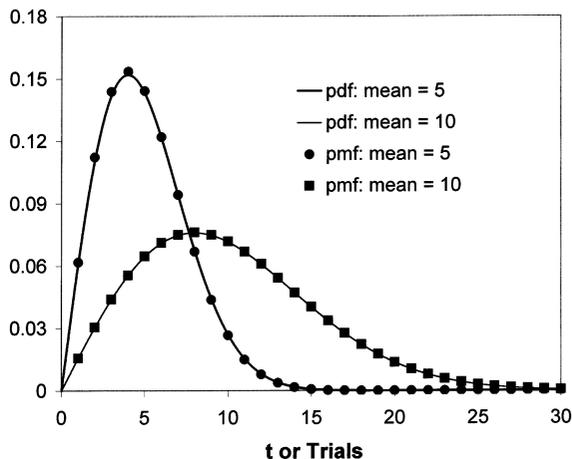


Fig. 2. Continuous and discrete versions of the Rayleigh distribution are compared as a function of continuous time (t) or discrete trials. The lines represent continuous probability density functions (pdf) with means of 5 and 10. The individual data points were generated by discrete probability mass functions (pmf) with the same means. The values of parameter β were 0.93815 and 0.98438.

discrete Rayleigh is difficult to work with. For example, it is difficult to obtain simple expressions for the mean and S.D. of the distribution. However, we may calculate these values numerically. Note in the continuous case, the coefficient of variation is constant, equal to $\sqrt{(4 - \pi)/\pi}$ as it should. It is not constant in the discrete case, but it is nearly so. From the derivation of the discrete case from the continuous case, we expect the value of e^{-m} to approximate that of β . Observe $e^{-\pi/50} = 0.93910 \approx 0.93815 = \beta$ in the case with a mean of 5, and $e^{-\pi/200} = 0.98441 \approx 0.98438 = \beta$ in the case with a mean of 10. Because the coefficient of variation is nearly constant, we may say that the discrete Rayleigh distribution ‘approximately’ has the scalar property.

7. Examples of the Rayleigh distribution

Many processes could generate the Rayleigh distribution. Egan (1975) describes an interesting example of the Rayleigh density function applied to signal detection theory:

Table 1

Comparison of descriptive statistics for continuous and corresponding discrete Rayleigh distributions with parameters chosen to match their means

Type	Parameter	Mean (μ)	S.D. (σ)	Coefficient of variation (σ/μ)
Discrete	$\beta = 0.93815$	4.9999	2.5811	0.51623
Continuous	$m = \pi/50$	5.0000	2.6136	0.52272
Discrete	$\beta = 0.98438$	10.0085	5.2156	0.52111
Continuous	$m = \pi/200$	10.0000	5.2272	0.52272

The Rayleigh density describes the following particular problem of a random walk. On each trial, a particle starts at the origin and takes two jumps, or steps. First, it takes a step of distance a along the horizontal axis; then, it takes a second step of distance b parallel to the vertical axis. Let a and b be independent, normally distributed variables; that is, $a, b: N(0, \sigma^2)$, and $E(ab) = 0$. After each trial, the distance from the origin to the particle is equal to $A = (a^2 + b^2)^{1/2}$, and the value of A has a Rayleigh distribution over trials. Given the distance from the origin, a power-law ROC results when an observer must decide from which one of two normal density functions, $N(0, \sigma_1^2)$ or $N(0, \sigma_2^2)$, the a and b were independently sampled... In ballistics, the Rayleigh density represents the distribution of the distance from the bullseye of the point of impact of a projectile that is ‘zeroed in.’ The projectile is acted upon by so many small, independent forces that the horizontal and vertical errors are each independently normally distributed. A power-law ROC results when an observer must decide which of two cannons was responsible for the magnitude of the ‘miss’ of the target (p. 92).

The interested reader may find additional examples of the Rayleigh distribution in psychophysical tasks involving the discrimination of small differences in the amplitudes of two sounds (Ronkin, 1969; Hautus and Irwin, 1992). For our purposes, it would be desirable to identify a neurophysiological pacemaker process that directly

produces the Rayleigh distribution in animal timing. Unfortunately, we are not prepared to propose a biological process that could be responsible.

8. Other distributions

The Rayleigh distribution is a special case of a Weibull distribution. The form of the distribution function for a Weibull is $F(t) = 1 - \exp(-(\rho t)^\kappa)$, with a failure rate function $\lambda(t) = \kappa\rho(\rho t)^{\kappa-1}$. Essentially, the Rayleigh distribution is the special case of the Weibull distribution with $\kappa = 2$, so that the power on t in the failure rate function $\lambda(t)$ is equal to unity. Ross (1984) indicates the Weibull distribution provides a close approximation to the distribution of a lifetime in the case of a ‘weakest link’ model; that is, if a system is composed of many components and failure occurs when any of the components fails. If one holds κ fixed, and only varies the parameter ρ in the Weibull family, then the resulting family of distributions will have a constant coefficient of variation; that is, they will have the scalar property. By varying κ , the coefficient of variation can be manipulated in the Weibull distribution. Because the Weibull distribution has one more parameter than does the Rayleigh distribution, it is not our preference as a pacemaker. Nevertheless, its interpretation as a weakest link model has some appeal. One could easily envisage a population of neurons in a pacemaker whose firing pattern depended on the lowest firing threshold (weakest link) of one of its components.

9. Summary and conclusions

A continuous or discrete version of a linear failure rate function assumes a linearly increasing value with increasing interval time. This result is equivalent to, but more parsimonious than, the accumulation of pulses in a Poisson-based pacemaker/switch/accumulator system. The linear failure rate functions define a family of Rayleigh distributions that produce a constant coefficient of variation independent of the slope parameter of the failure rate function. Therefore, they always produce the scalar property. This parsimonious pacemaker may be incorporated into existing models of timing. Because it predicts the scalar property even before its value is compared with the reference memory, its utilization may eliminate the need for additional sources of variance in order to produce the scalar property, or the requirement that comparison to the reference memory must be based on a ratio process.

Our model has an interesting implication for an important difference between SET and the behavioral theory of timing (BeT) (Killeen and Fetterman, 1988). SET assumes that the pacemaker rate is fixed and independent of the reinforcement rate, whereas BeT assumes that the pacemaker rate is proportional to the reinforcement rate. The assumption of a fixed pacemaker rate is the equivalent of assuming m is a constant in our linear failure rate function, $\lambda(t) = mt$. If m is allowed to vary with the reinforcement rate, then the pacemaker rate may be proportional to the reinforcement rate. Thus, the interpretation of this single parameter may allow, or may preclude, the pace-

maker rate being influenced by the reinforcement rate.

References

- Balakrishnan, J.D., 1994. Simple additivity of stochastic psychological processes: tests and measures. *Psychometrika* 59 (2), 217–240.
- Cox, D.R., Oakes, D., 1984. *Analysis of Survival Data*. Chapman and Hall, New York.
- Egan, J.P., 1975. *Signal Detection Theory and ROC Analysis*. Academic Press, New York.
- Gibbon, J., 1977. Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84 (3), 279–325.
- Gibbon, J., 1991. Origins of scalar timing. *Learn. Motiv.* 22, 3–38.
- Gibbon, J., Church, R.M., 1984. Sources of variance in an information processing theory of timing. In: Roitblat, H.L., Bever, T.G., Terrace, H.S. (Eds.), *Animal Cognition*. Lawrence Erlbaum, Hillsdale, NJ.
- Hautus, M.J., Irwin, R.J., 1992. Amplitude discrimination of sinusoids and narrow-band noise with Rayleigh properties. *Percept. Psychophys.* 52 (1), 53–62.
- Killeen, P.R., Fetterman, J.G., 1988. A behavioral theory of timing. *Psychol. Rev.* 95, 274–295.
- Luce, R.D., 1986. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, New York.
- Mechner, F., 1958. Probability relations within response sequences under ratio reinforcement. *J. Exp. Anal. Behav.* 1, 109–1121.
- Ronkin, D.A., 1969. Intensity discrimination of Rayleigh noise. *J. Acoust. Soc. Am.* 45, 54–57.
- Ross, S., 1980. *Introduction to Probability Models*, 2nd ed. Academic Press, New York.
- Ross, S., 1984. *A First Course in Probability*, 2nd ed. Macmillan, New York.
- Townsend, J.T., Ashby, F.G., 1983. *Stochastic Modeling of Elementary Psychological Processes*. Cambridge University Press, Cambridge.