The Poisson Process and Associated Probability Distributions on Time Scales

Dylan R. Poulsen
Department of Mathematics
Baylor University
Waco, TX 76798
Email: Dylan_Poulsen@baylor.edu

Michael Z. Spivey
Department of Mathematics and
Computer Science
University of Puget Sound
Tacoma, WA 98416
Email: mspivey@pugetsound.edu

Robert J. Marks II
Department of Electrical and
Computer Engineering
Baylor University
Waco, TX 76798
Email: Robert_Marks@baylor.edu

Abstract—**Duals of probability distributions on continuous ($\mathbb{R}$) domains exist on discrete ($\mathbb{Z}$) domains. The Poisson distribution on $\mathbb{R}$, for example, manifests itself as a binomial distribution on $\mathbb{Z}$. Time scales are a domain generalization in which $\mathbb{R}$ and $\mathbb{Z}$ are special cases. We formulate a generalized Poisson process on an arbitrary time scale and show that the conventional Poisson distribution on $\mathbb{R}$ and binomial distribution on $\mathbb{Z}$ are special cases. The waiting times of the generalized Poisson process are used to derive the Erlang distribution on a time scale and, in particular, the exponential distribution on a time scale. The memoryless property of the exponential distribution on $\mathbb{R}$ is well known. We find conditions on the time scale which preserve the memorylessness property in the generalized case.**

I. INTRODUCTION

The theory of continuous and discrete time stochastic processes is well developed [7], [8]. Stochastic processes on general closed subsets of the real numbers, also known as *time scales*, allow a generalization to other domains [4], [9]. The notion of a stochastic process on time scales naturally leads to questions about probability theory on time scales, which has been developed by Kahraman [5]. We begin by introducing a generalized Poisson process on time scales and show it reduces to the conventional Poisson process on $\mathbb{R}$ and the binomial distribution on $\mathbb{Z}$. We then use properties of the Poisson process to motivate generalized Erlang and exponential distributions on time scales. Finally, we show that the generalized exponential distribution has an analogue of the memorylessness property under periodicity conditions on the time scale.

II. FOUNDATIONS

A time scale, $\mathbb{T}$, is any closed subset of the real line. We restrict attention to causal time scales [6] where $0 \in \mathbb{T}$ and $t \geq 0$ for all $t \in \mathbb{T}$. The *forward jump operator* [2], [10], $\sigma(t)$, is defined as the point immediately to the right of t, in the sense that $\sigma(t) = \inf\{s \in \mathbb{T} \,\forall\, s > t\}$. The *graininess* is the distance between points defined as $\mu(t) := \sigma(t) - t$. For $\mathbb{R}$, $\sigma(t) = t$ and $\mu(t) = 0$.

The time scale or *Hilger derivative* of a function $x(t)$ on $\mathbb{T}$ is defined as

$$x^{\Delta}(t) := \frac{x(\sigma(t)) - x(t)}{\mu(t)}. \qquad (\text{II.1})$$

On $\mathbb{R}$, this is interpreted in the limiting case and $x^{\Delta}(t) = \frac{d}{dt} x(t)$. The *Hilger integral* can be viewed as the antiderivative in the sense that, if $y(t) = x^{\Delta}(t)$, then for $s, t \in \mathbb{T}$,

$$\int_{\tau=s}^{t} y(\tau)\Delta\tau = x(t) - x(s).$$

The solution to the differential equation

$$x^{\Delta}(t) = zx(t); \; x(0) = 1,$$

is $x(t) = e_z(t, 0)$ where [2], [10]

$$e_z(t, s) := \exp\left(\int_{\tau=s}^{t} \frac{\text{Log}(1 + \mu(\tau)z)}{\mu(\tau)}\Delta\tau\right).$$

For an introduction to time scales, there is an online tutorial [10] or, for a more thorough treatment, see the text by Bohner and Peterson [2].

III. THE POISSON PROCESS ON TIME SCALES

We begin by presenting the derivation for a particular stochastic process on time scales which mirrors a derivation for the Poisson process on $\mathbb{R}$ [3].

Let $\lambda > 0$. Assume the probability an event occurs in the interval $[t, \sigma(s))_{\mathbb{T}}$ is given by

$$-(\ominus\lambda)(t)(\sigma(s) - t) + \text{o}(s - t).$$

where $\ominus z := -z/(1 - \mu(t)z)$ [2], [10]. Hence the probability that no event occurs on the interval is given by

$$1 + (\ominus\lambda)(t)(\sigma(s) - t) + \text{o}(s - t).$$

We also assume that at $t = 0$ no events have occurred.

We now define a useful notation. Let $X : \mathbb{T} \to \mathbb{N}^0$ be a counting process [8] where $\mathbb{N}^0$ denotes all nonnegative integers. For $k \in \mathbb{N}^0$, define $p_k(t) = \mathbb{P}[X(t) = k]$, the probability that k events have occurred by time $t \in \mathbb{T}$. Let $t, s \in \mathbb{T}$ with $s > t$. Consider the successive intervals $[0, t)_{\mathbb{T}}$

and $[t, \sigma(s))_{\mathbb{T}}$. We can therefore set up the system of equations

$$
\begin{aligned}
p_0(\sigma(s)) &= p_0(t)[1 + (\ominus\lambda)(t)(\sigma(s) - t)] + o(s - t) \\
p_1(\sigma(s)) &= p_1(t)[1 + (\ominus\lambda)(t)(\sigma(s) - t)] \\
&+ p_0(t)[-(\ominus\lambda)(t)(\sigma(s) - t)] + o(s - t) \\
&\vdots \\
p_k(\sigma(s)) &= p_k(t)[1 + (\ominus\lambda)(t)(\sigma(s) - t)] \\
&+ p_{k-1}(t)[-(\ominus\lambda)(t)(\sigma(s) - t)] + o(s - t) \\
&\vdots
\end{aligned}
$$

with initial conditions $p_0(0) = 1$ and $p_k(0) = 0$ for $k > 0$. We will let $s \to t$ and solve these equations recursively. Consider the p_0 equation. By the definition of the derivative on time scales, we have

$$
p_0^\Delta(t) = \lim_{s \to t} \frac{p_0(\sigma(s)) - p_0(t)}{\sigma(s) - t} = (\ominus\lambda)(t)p_0(t),
$$

which, using the initial value $p_0(0) = 1$, has a solution

$$
p_0(t) = e_{\ominus\lambda}(t, 0). \tag{III.1}
$$

Now consider the p_1 equation. Substituting the solution of the p_0 equation yields

$$
\begin{aligned}
p_1(\sigma(s)) &= p_1(t)[1 + (\ominus\lambda)(t)(\sigma(s) - t)] \\
&+ e_{\ominus\lambda}(t, 0)[-(\ominus\lambda)(t)(\sigma(s) - t)] + o(s - t),
\end{aligned}
$$

which, using (II.1), yields

$$
p_1^\Delta(t) = (\ominus\lambda)(t)p_1(t) - (\ominus\lambda)(t)e_{\ominus\lambda}(t, 0). \tag{III.2}
$$

Using the variation of constants formula on time scales [2], we arrive at the solution

$$
\begin{aligned}
p_1(t) &= -\int_0^t e_{\ominus\lambda}(t, \sigma(\tau))(\ominus\lambda)(\tau)e_{\ominus\lambda}(\tau, 0)\Delta\tau \\
&= -\int_0^t e_\lambda(\tau, t)(1 + \mu(\tau)\lambda)(\ominus\lambda)(\tau)e_{\ominus\lambda}(\tau, 0)\Delta\tau \\
&= \lambda\int_0^t e_\lambda(\tau, 0)e_\lambda(0, t)e_{\ominus\lambda}(\tau, 0)\Delta\tau \\
&= \lambda\int_0^t e_{\ominus\lambda}(t, 0)\Delta\tau \\
&= \lambda t e_{\ominus\lambda}(t, 0) \\
&= \frac{\lambda}{1 + \mu(0)\lambda}t e_{\ominus\lambda}(t, \sigma(0)) \\
&= -(\ominus\lambda)(0)t e_{\ominus\lambda}(t, \sigma(0)).
\end{aligned}
$$

Now consider the p_2 equation. Substituting the solution of the p_1 equation yields

$$
\begin{aligned}
p_2(\sigma(s)) &= p_2(t)[1 + (\ominus\lambda)(t)(\sigma(s) - t)] \\
&- (\ominus\lambda)(0)t e_{\ominus\lambda}(t, \sigma(0))[-(\ominus\lambda)(t)(\sigma(s) - t)] \\
&+ o(s - t),
\end{aligned}
$$

which, using (II.1) yields

$$
p_2^\Delta(t) = (\ominus\lambda)(t)p_2(t) + (\ominus\lambda)(0)(\ominus\lambda)(t)t e_{\ominus\lambda}(t, \sigma(0)).
$$

Again, using the variation of constants formula on time scales, we arrive at the solution

$$
\begin{aligned}
p_2(t) &= (\ominus\lambda)(0) \\
&\times \int_0^t e_{\ominus\lambda}(t, \sigma(\tau))(\ominus\lambda)(\tau)\tau e_{\ominus\lambda}(\tau, \sigma(0))\Delta\tau \\
&= (\ominus\lambda)(0) \\
&\times \int_0^t e_\lambda(\tau, t)(1 + \mu(\tau)\lambda)(\ominus\lambda)(\tau)\tau e_{\ominus\lambda}(\tau, \sigma(0))\Delta\tau \\
&= -\lambda(\ominus\lambda)(0) \\
&\times \int_0^t \tau e_\lambda(\tau, \sigma(0))e_\lambda(\sigma(0), t)e_{\ominus\lambda}(\tau, \sigma(0))\Delta\tau \\
&= -\lambda(\ominus\lambda)(0)e_{\ominus\lambda}(t, \sigma(0))\int_0^t \tau\Delta\tau \\
&= -\lambda(\ominus\lambda)(0)e_{\ominus\lambda}(t, \sigma(0))h_2(t, 0) \\
&= \frac{-\lambda}{1 + \mu(\sigma(0))\lambda}(\ominus\lambda)(0)e_{\ominus\lambda}(t, \sigma^2(0))h_2(t, 0) \\
&= (\ominus\lambda)(\sigma(0))(\ominus\lambda)(0)h_2(t, 0)e_{\ominus\lambda}(t, \sigma^2(0)).
\end{aligned}
$$

In general, it can be shown via induction that

$$
p_k(t) = (-1)^k h_k(t, 0)e_{\ominus\lambda}(t, \sigma^k(0))\prod_{i=0}^{k-1}(\ominus\lambda)(\sigma^i(0)),
$$

where $h_k(t, 0)$ is the kth generalized Taylor monomial [2].

The above derivation motivates the following definition:

Definition III.1. Let $\mathbb{T}$ be a time scale. We say $S : \mathbb{T} \to \mathbb{N}^0$ is a $\mathbb{T}$–*Poisson process* with rate $\lambda > 0$ if for $t \in \mathbb{T}$ and $k \in \mathbb{N}^0$,

$$
\mathbb{P}[S(t; \lambda) = k] = (-1)^k h_k(t, 0)e_{\ominus\lambda}(t, \sigma^k(0))\prod_{i=0}^{k-1}(\ominus\lambda)(\sigma^i(0)). \tag{III.3}
$$

Each fixed $t \in \mathbb{T}$ generates a discrete distribution of the number of arrivals at t. We now examine the specific examples of $\mathbb{R}$, $\mathbb{Z}$ and the harmonic time scale [2].

A. On $\mathbb{R}$ and $\mathbb{Z}$

Let $S : \mathbb{R} \to \mathbb{N}^0$ be an $\mathbb{R}$–Poisson process. Then $\sigma^i(0) = 0$ for all $i \in \mathbb{N}$, $(\ominus\lambda)(t) = -\lambda$ for all $t \in \mathbb{R}$ and $h_k(t) = \frac{t^k}{k!}$. Thus we have

$$
\mathbb{P}[S(t; \lambda) = k] = \frac{(\lambda t)^k}{k!}e^{-\lambda t},
$$

which we recognize as the Poisson distribution.

Now let $S : \mathbb{Z} \to \mathbb{N}^0$ be an $\mathbb{N}^0$–Poisson process. We have $\sigma^i(0) = i$ for all $i \in \mathbb{N}$, $(\ominus\lambda)(t) = \frac{-\lambda}{1+\lambda} := -p$, and $h_k(t) = \binom{t}{k}$. Thus we have

$$
\mathbb{P}[S(t; \lambda) = k] = \binom{t}{k}p^k(1-p)^{t-k},
$$

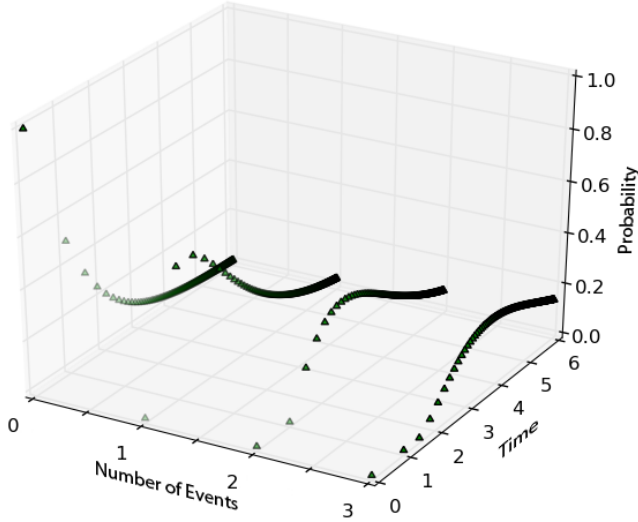which we recognize as the binomial distribution.

Fig. 1. Probability against *Number of Events* and *Time* for the $\mathbb{H}_n$–Poisson Process with rate 1.
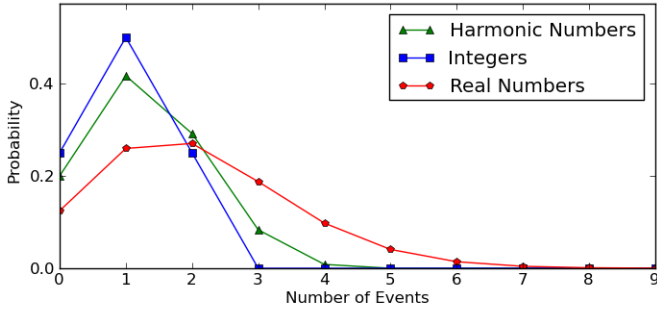

Fig. 2. A comparison of probability versus number of events near $t = 2$ for the $\mathbb{H}_n$–Poisson process with rate 1, the $\mathbb{R}$–Poisson process with rate 1 and the $\mathbb{Z}$–Poisson process with rate 1. Note that the $\mathbb{H}_n$–Poisson process behaves more like the $\mathbb{Z}$–Poisson process than the $\mathbb{R}$–Poisson process.

B. On the Harmonic Time Scale

Now let $S : \mathbb{H}_n \to \mathbb{N}^0$ be an $\mathbb{H}_n$–Poisson process with rate λ, where

$$t \in \mathbb{H}_n \text{ if and only if } t = \sum_{k=1}^{n} \frac{1}{k} \text{ for some } n \in \mathbb{N},$$

which we call the harmonic time scale. To help understand later figures and emphasize that S yields a distinct discrete distribution for each value of t, we show the probability against the number of events and time in Figure 1. The choice of $\mathbb{H}_n$ as the time scale shows very informative behavior. Near $t = 0$, when the graininess is large, we find behavior that is more like the integers. In contrast, away from $t = 0$, where the graininess is small, we find behavior that is more like the real numbers. This behavior is demonstrated in Figures 2–4.
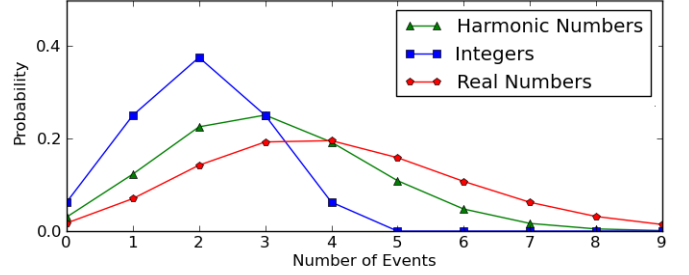

Fig. 3. A comparison of probability versus number of events near $t = 4$ for the $\mathbb{H}_n$–Poisson process with rate 1, the $\mathbb{R}$–Poisson process with rate 1 and the $\mathbb{Z}$–Poisson process with rate 1. Note that the $\mathbb{H}_n$–Poisson process behaves more like the $\mathbb{R}$–Poisson process than the $\mathbb{Z}$–Poisson process.
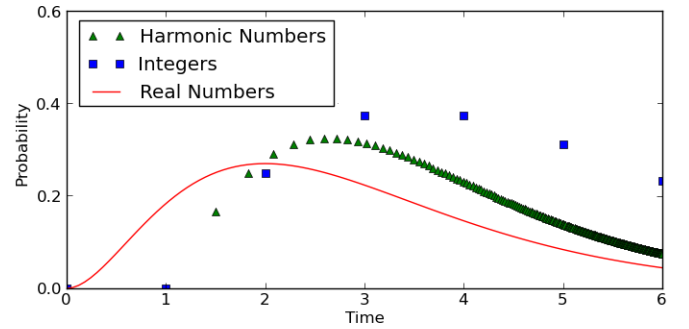

Fig. 4. A comparison of probability versus time when we fix the number of events at 2 for the $\mathbb{H}_n$–Poisson process with rate 1, the $\mathbb{R}$–Poisson process with rate 1 and the $\mathbb{Z}$–Poisson process with rate 1. Note that the $\mathbb{H}_n$–Poisson process behaves more like the $\mathbb{Z}$–Poisson process near $t = 0$ and more like the $\mathbb{R}$–Poisson process away from $t = 0$.

IV. THE ERLANG DISTRIBUTION ON TIME SCALES

A time scales generalization of the Erlang distribution can be generated by examining the waiting times between any number of events in the $\mathbb{T}$–Poisson process. To that end, let $\mathbb{T}$ be a time scale. Let $S : \mathbb{T} \to \mathbb{N}$ be a $\mathbb{T}$–Poisson process with rate λ. Let T_n be a random variable which denotes the time until the nth event. We have

$$\begin{aligned} \mathbb{P}[S(t;\lambda) < n] &= \mathbb{P}[T_n > t] \\ &= 1 - \mathbb{P}[T_n \leq t]. \end{aligned}$$

which implies

$$1 - \sum_{k=0}^{n-1} \mathbb{P}[S(t;\lambda) = k] = \mathbb{P}[T_n \leq t],$$

which motivates the following definition.

Definition IV.1. Let $\mathbb{T}$ be a time scale, $S : \mathbb{T} \to \mathbb{N}^0$ be a $\mathbb{T}$-Poisson Process with rate $\lambda > 0$. We say $F(t; n, \lambda)$ is the $\mathbb{T}$–*Erlang cumulative distribution function* with shape parameter

n and rate λ provided

$$F(t;n,\lambda) = 1 - \sum_{k=0}^{n-1} \mathbb{P}[S(t;\lambda) = k].$$

From our derivation, it is clear that the $\mathbb{T}$–Erlang distribution models the time until the nth event in the $\mathbb{T}$–Poisson process. We would like to know the probability that the nth event is in any subset of $\mathbb{T}$. To this end, we introduce the $\mathbb{T}$–Erlang probability density function in the next definition.

Definition IV.2. Let $\mathbb{T}$ be a time scale, $S : \mathbb{T} \to \mathbb{N}^0$ be a $\mathbb{T}$-Poisson Process with rate $\lambda > 0$. We say $f(t;n,\lambda)$ is the $\mathbb{T}$–*Erlang probability density function* with shape parameter n and rate λ provided

$$f(t;n,\lambda) = -\sum_{k=0}^{n-1} [\mathbb{P}[S(t;\lambda) = k]]^\Delta.$$

where the Δ–differentiation is with respect to t.

We want to show that $f(t;n,\lambda)$ can rightly be called a probability density with respect to some accumulation function. Thus, we have the following theorem.

Theorem IV.1. *Let $\mathbb{T}$ be a time scale. Let $F(t;n,\lambda)$ be a $\mathbb{T}$–Erlang cumulative distribution function with shape parameter n and rate λ and let $f(t;n,\lambda)$ be a $\mathbb{T}$–Erlang probability density function with shape parameter n and rate λ. Then*

$$\int_0^t f(\tau;n,\lambda)\Delta\tau = F(t;n,\lambda) \qquad \text{(IV.1)}$$

and in particular

$$\int_\mathbb{T} f(\tau;n,\lambda)\Delta\tau = 1. \qquad \text{(IV.2)}$$

Proof: Implicit in the definition of the $\mathbb{T}$–Erlang probability distribution is a $\mathbb{T}$-Poisson process $S : \mathbb{T} \to \mathbb{N}^0$. By the assumption that

$$\mathbb{P}[S(0;\lambda) = k] = \begin{cases} 1 & k = 0 \\ 0 & k > 0, \end{cases}$$

we have,

$$
\begin{aligned}
\int_0^t f(\tau;n,\lambda)\Delta\tau &= \int_0^t -\sum_{k=0}^{n-1} \mathbb{P}[S(\tau;\lambda) = k]^\Delta \Delta\tau \\
&= -\sum_{k=0}^{n-1} \int_0^t \mathbb{P}[S(\tau;\lambda) = k]^\Delta \Delta\tau \\
&= -\sum_{k=0}^{n-1} \mathbb{P}[S(\tau;\lambda) = k]|_0^t \\
&= -\sum_{k=0}^{n-1} \mathbb{P}[S(t;\lambda) = k] \\
&\quad + \sum_{k=0}^{n-1} \mathbb{P}[S(0;\lambda) = k] \\
&= 1 - \sum_{k=0}^{n-1} \mathbb{P}[S(t;\lambda) = k] \\
&= F(t;n,\lambda),
\end{aligned}
$$

which proves (IV.1). To prove (IV.2), we note for all $k < n$,

$$\lim_{t\to\infty} \mathbb{P}[S(t;\lambda) = k] = 0,$$

by repeated application of L'Hôpital's rule for time scales on III.3 [1]. This fact proves (IV.2) by the same argument as the proof of (IV.1). ∎

We note that the moments of the $\mathbb{T}$–Erlang distribution cannot in general be calculated explicitly without some knowledge of the time scale.

V. THE EXPONENTIAL DISTRIBUTION ON TIME SCALES

Of particular interest to us is the $\mathbb{T}$–Erlang distribution with shape parameter 1. By the above discussion and equation (III.1), the probability density function of this distribution is given by

$$f(t;1,\lambda) = -\mathbb{P}[S^\Delta(t;\lambda) = 0] = -(\ominus\lambda)(t)e_{\ominus\lambda}(t,0).$$

Definition V.1. Let $\mathbb{T}$ be a time scale and let T be a $\mathbb{T}$–Erlang random variable with shape parameter 1 and rate λ. Then we say T is a $\mathbb{T}$–*exponential random variable* with rate λ.

A. The Expected Value

The $\mathbb{T}$–exponential distribution gives us the rare opportunity to calculate a moment without any knowledge of the time scale.

Lemma V.1. *Let $\mathbb{T}$ be a time scale and let T be a $\mathbb{T}$–exponential random variable with rate $\lambda > 0$. Then*

$$\mathbb{E}(T) = \frac{1}{\lambda}.$$

Proof: Using integration by parts on time scales, we find

$$
\begin{aligned}
\mathbb{E}(T) &= \int_0^\infty t[-(\ominus\lambda)(t)e_{\ominus\lambda}(t,0)]\Delta t \\
&= -te_{\ominus\lambda}(t,0)\big|_0^\infty + \int_0^\infty e_{\ominus\lambda}(\sigma(t),0)\Delta t \\
&= 0 + \int_0^\infty (1+\mu(t)(\ominus\lambda)(t))e_{\ominus\lambda}(t,0)\Delta t \\
&= \int_0^\infty \frac{1}{1+\mu(t)\lambda}e_{\ominus\lambda}(t,0)\Delta t \\
&= -\frac{1}{\lambda}\int_0^\infty \frac{-\lambda}{1+\mu(t)\lambda}e_{\ominus\lambda}(t,0)\Delta t \\
&= -\frac{1}{\lambda}\int_0^\infty (\ominus\lambda)(t)e_{\ominus\lambda}(t,0)\Delta t \\
&= -\frac{1}{\lambda}e_{\ominus\lambda}(t,0)\big|_0^\infty \\
&= -\frac{1}{\lambda}[0-1] \\
&= \frac{1}{\lambda},
\end{aligned}
$$

which proves our claim. ∎

B. On $\mathbb{R}$ and $\mathbb{Z}$

We note that if $\mathbb{T} = \mathbb{R}$, then we have

$$ f(t;1,\lambda) = \lambda e^{-\lambda t}, $$

which we recognize as the exponential distribution. By Lemma V.1, we find the mean of the exponential distribution is $1/\lambda$, which is well known.

Now if $\mathbb{T} = \mathbb{Z}$, then we have

$$ f(t;1,\lambda) = \frac{\lambda}{1+\lambda}\left(1-\frac{\lambda}{1+\lambda}\right)^t = p(1-p)^t, $$

where $p := \frac{\lambda}{1+\lambda}$. We recognize the above as the geometric distribution. By Lemma V.1, we find the mean of the geometric distribution is $1/\lambda = (1-p)/p$.

C. The ω–Memorylessness Property

Both the geometric and exponential distributions are completely classified by the fact that they have the memorylessness property [8]. We recall the the memoryless property on $\mathbb{R}$ is the property that if T is a continuous random variable, then for all $t, \tau \in \mathbb{R}$,

$$ \mathbb{P}[T > t+\tau | T > t] = \mathbb{P}[T > \tau] $$

and that the memoryless property on $\mathbb{Z}$ is the property that if T is a discrete random variable, then for all $t, \tau \in \mathbb{Z}$,

$$ \mathbb{P}[T > t+\tau | T > t] = \mathbb{P}[T > \tau]. $$

We would like to find conditions on the time scale $\mathbb{T}$ such that the $\mathbb{T}$–exponential distribution on time scales has this property. Let $\mathbb{T}$ is ω-periodic, that is, if $t \in \mathbb{T}$ then $t+\omega \in \mathbb{T}$. Then we can define a property much like the memorylessness property.

Definition V.2. Let $\mathbb{T}$ be an ω–periodic time scale. We say a probability distribution on $\mathbb{T}$ has the ω–*memorylessness property* provided for all $t \in \mathbb{T}$,

$$ P(T > t+\omega | T > t) = P(T > \omega), $$

We note that this definition generalizes the memorylessness property on $\mathbb{R}$ and $\mathbb{Z}$ since $\mathbb{R}$ and $\mathbb{Z}$ are ω–periodic for any ω in $\mathbb{R}$ and $\mathbb{Z}$, respectively.

Let $\mathbb{T}$ be ω–periodic and let T be a $\mathbb{T}$–exponential random variable. Then we claim the $\mathbb{T}$–exponential distribution has the ω–memorylessness property. To show this claim, we first prove two lemmas.

Lemma V.2. *Let $\mathbb{T}$ be an ω-periodic time scale and let $\lambda > 0$. Then for $t, t_0 \in \mathbb{T}$, $e_{\ominus\lambda}(t+\omega, t_0) = e_{\ominus\lambda}(t, t_0-\omega)$.*

Proof: By the definition of the time scales exponential function,

$$
\begin{aligned}
e_{\ominus\lambda}(t+\omega, t_0) &= \exp\left(\int_{t_0}^{t+\omega} \frac{\text{Log}(1+(\ominus\lambda)(s)\mu(s))\Delta s}{\mu(s)}\right) \\
&= \exp\left(\int_{t_0}^{t+\omega} \frac{\text{Log}\left(1+\frac{-\lambda\mu(s)}{1+\lambda\mu(s)}\right)\Delta s}{\mu(s)}\right) \\
&= \exp\left(\int_{t_0-\omega}^{t} \frac{\text{Log}\left(1+\frac{-\lambda\mu(\tau+\omega)}{1+\lambda\mu(\tau+\omega)}\right)\Delta\tau}{\mu(\tau+\omega)}\right) \\
&= \exp\left(\int_{t_0-\omega}^{t} \frac{\text{Log}\left(1+\frac{-\lambda\mu(\tau)}{1+\lambda\mu(\tau)}\right)\Delta\tau}{\mu(\tau)}\right) \\
&= \exp\left(\int_{t_0-\omega}^{t} \frac{\text{Log}(1+(\ominus\lambda)(\tau)\mu(\tau))\Delta\tau}{\mu(\tau)}\right) \\
&= e_{\ominus\lambda}(t, t_0-\omega),
\end{aligned}
$$

where we use the fact that for ω-periodic time scales $\mu(t+\omega) = \mu(t)$ for all $t \in \mathbb{T}$ and the change of variables $\tau = s-\omega$. ∎

Lemma V.3. *Let $\mathbb{T}$ be an ω-periodic time scale and $\lambda > 0$. Then for all $t \in \mathbb{T}$, $e_{\ominus\lambda}^\Delta(t+\omega, t) = 0$.*

Proof: By the product rule on time scales and Lemma V.2,

$$
\begin{aligned}
e_{\ominus\lambda}^\Delta(t+\omega, t) &= (e_{\ominus\lambda}(t+\omega, t_0)e_{\ominus\lambda}(t_0, t))^\Delta \\
&= (e_{\ominus\lambda}(t, t_0-\omega)e_{\ominus\lambda}(t_0, t))^\Delta \\
&= (e_{\ominus\lambda}(t, t_0-\omega)e_\lambda(t, t_0))^\Delta \\
&= e_{\ominus\lambda}(\sigma(t), t_0-\omega)\lambda e_\lambda(t, t_0) \\
&\quad + (\ominus\lambda)(t)e_{\ominus\lambda}(t, t_0-\omega)e_\lambda(t, t_0) \\
&= \lambda(1+(\ominus\lambda)(t)\mu(t))e_{\ominus\lambda}(t, t_0-\omega)e_\lambda(t, t_0) \\
&\quad + (\ominus\lambda)(t)e_{\ominus\lambda}(t, t_0-\omega)e_\lambda(t, t_0) \\
&= [-(\ominus\lambda)(t)+(\ominus\lambda)(t)] \\
&\quad e_{\ominus\lambda}(t, t_0-\omega)e_\lambda(t, t_0) \\
&= 0.
\end{aligned}
$$

∎

The above lemmas allow us to prove the following result.

Theorem V.4. *Let $\mathbb{T}$ be an ω–periodic time scale and let $\lambda > 0$. Then the $\mathbb{T}$–exponential distribution with rate λ has the ω–memorylessness property.*

Proof: Let T be a $\mathbb{T}$–exponential random variable with rate $\lambda > 0$. By Lemma V.2 and Lemma V.3,

$$
\begin{aligned}
P(T > t + \omega | T > t) &= \frac{P(T > t + \omega)}{P(T > t)} \\
&= \frac{\int_{t+\omega}^{\infty} -(\ominus \lambda)(\tau) e_{\ominus \lambda}(\tau, 0) \Delta \tau}{\int_{t}^{\infty} -(\ominus \lambda)(\tau) e_{\ominus \lambda}(\tau, 0) \Delta \tau} \\
&= \frac{e_{\ominus \lambda}(t + \omega, 0)}{e_{\ominus \lambda}(t, 0)} \\
&= e_{\ominus \lambda}(t + \omega, t) \\
&= e_{\ominus \lambda}(\omega, 0) \\
&= P(T > \omega),
\end{aligned}
$$

since $e_{\ominus \lambda}(\omega, 0)$ is a constant independent of t by Lemma V.3. Thus the $\mathbb{T}$–exponential distribution has the ω–memorylessness property ∎

REFERENCES

[1] M. Bohner and A. Peterson, *Advances in Dynamic Equations on Time Scales*, Birkhäuser, Boston, 2003.

[2] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales*, Birkhäuser, Boston, 2001.

[3] W. Ching and M. Ng, *Markov chains: models, algorithms and applications*, Springer, New York, 2006.

[4] John M. Davis, Ian A. Gravagne and Robert J. Marks II, "Bilateral Laplace Transforms on Time Scales: Convergence, Convolution, and the Characterization of Stationary Stochastic Time Series," Circuits, Systems, and Signal Processing, Birkhäuser, Boston, Volume 29, Issue 6 (2010), Page 1141. [DOI 10.1007/s00034-010-9196-2]

[5] S. Kahraman, "Probability Theory Applications on Time Scales," M.S. Thesis, İzmir Institute of Technology, 2008

[6] Robert J. Marks II, Ian A. Gravagne and John M. Davis, "A Generalized Fourier Transform and Convolution on Time Scales," Journal of Mathematical Analysis and Applications Volume 340, Issue 2, 15 April 2008, Pages 901-919.

[7] R.J. Marks II, *Handbook of Fourier Analysis and Its Applications,*" Oxford University Press (2009).

[8] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd Edition, McGraw-Hill, New York (1991)

[9] S. Sanyal, "Stochastic Dynamic Equations," Ph.D. Thesis, Missouri University of Science and Technology, 2008

[10] Baylor Time Scales Group, http://timescales.org/