Physica A 643 (2024) 129827

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Binomial vs. Poisson statistics: From a toy model to a stochastic model for radioactive decay

Sergio Sánchez-Sánchez [a,*], Ernesto Cortés-Pérez [b], Víctor I. Moreno-Oliva [a]

[a] *Department of Applied Mathematics, Universidad del Istmo, Ciudad Universitaria S/N, Barrio Santa Cruz, 4a. Sección Tehuantepec, Oaxaca, México Zip Code 70760, Mexico*

[b] *Department of Computer Engineering, Universidad del Istmo, Ciudad Universitaria S/N, Barrio Santa Cruz, 4a. Sección Tehuantepec, Oaxaca, México Zip Code 70760, Mexico*

ARTICLE INFO

Keywords:
Radioactive decay
Stochastic process
Binomial and Poisson statistics
Stochastic entropy

ABSTRACT

Disintegration is a physical phenomenon of atomic nuclei —radioactive isotopes decay— has been modeled with different approaches (deterministic and random), from didactic *toy models* using as reference the roll (experiment) of standard six-sided dice (Arthur and Ian, 2012), to the generalization of probabilistic methods. Using for example the moment generating function (**MGF**) method, to obtain the behavior of its probability distribution for n multi-sided dice, —i.e., dice of $s \geq 6$ sided (Singh et al., 2011, Sánchez-Sánchez et al., 2022)—. Radioactive decay is essentially statistical (random) in nature, so we cannot predict when any of the atoms will decay. The MGF method and stochastic models were applied to the so-called radioactive dice (*toy model*), to obtain a theoretical Poisson-like distribution (*stochastic model*). In this work, we carry out an exhaustive study and a comparison —on the apparent *discrepancy*— of the Binomial and Poisson statistics associated with the decay of radioactive nuclei. To gain a deeper understanding of this phenomenon —radioactive decay— we use the theory of stochastic processes, i.e., modeling these distributions in a stochastic context using the efficient mathematical tools of this theory. They are discrete random variable processes in continuous time. So, we use an approach of the so-called master equation (Kolmogorov equations). We study them in general as *Birth-Death processes* —both processes separately, highlighting their disagreements— modeling the Poisson process as a *Pure Death process*. We solve the master equations of the Poisson process by introducing the so-called *sojourn time*. Also, we study the relative fluctuations through the *Fano factor*. We analyze the deeper concept of *Entropy* of the binomial and Poisson processes by calculating their metrics —In *Shannon's information theory* context—. We show that they have a corresponding statistical link between both images and the radioactive decay distributions. In this way, we gain a deeper insight into the random nature of nuclear decay with its stochastic distributions. In summary, this paper addresses the extension of deterministic systems that are, in reality, of a random nature within a theoretical framework of so-called stochastic processes and information theory (entropy). We link entropy stochastic (binomial and Poisson) and its intrinsic fluctuations with the physical mechanisms of the collective dynamics of radioactive decay.

* Corresponding author.
E-mail address: ssanys1@hotmail.com (S. Sánchez-Sánchez).

https://doi.org/10.1016/j.physa.2024.129827
Received 9 August 2022; Received in revised form 29 April 2024
Available online 14 May 2024

1. Introduction

It is harder to crack prejudice than an atom. A. Einstein.

Non-deterministic events —*random or stochastic*— occur frequently in natural phenomena and physical systems, necessitating the development of statistical-probabilistic models to study their evolution. These models help us understand and forecast the behavior of these events over time, known as stochastic processes. Some examples include weather patterns, earthquakes, stock market predictions, radioactive decay in atoms, nuclear decay, and even outcomes in casino gambling [1]. Hence, the research of stochastic processes is currently a crucial mathematical and conceptual tool in physics, information sciences, engineering, and technological applications. Both classical and quantum physics have delved into stochastic modeling to comprehend systems in which, owing to their random nature, it is essential to regard variables as inherently random [2,3]. For example, a case of particular interest is the so-called radioactive dice system [4–7]. Despite being a didactic tool in high school physics and chemistry labs and classrooms, it aids in comprehending the behavior of nuclei —*radioisotopes*— in radioactive activity. Murray and Hart [4] demonstrate this application on a didactic level, enabling analysis of the trend of radioactive dice as the number of dice and sides increases. According to this study, the output of the radioactive dice will adhere to a *deterministic exponential decay law*.

On the other hand, Singh, A.K. et al. in Ref. [8], utilize the standard Moment Generating Function (**MGF**) method to investigate the probabilities of unbiased dice throwing. They apply this method to gambling games in casinos (such as those in Las Vegas), employing standard six-sided dice to predict the odds associated with each simultaneous roll of 1 to 5 dice. This method is further generalized by Sánchez et al. [9] for multi-sided dice, with consideration given to its application to the so-called radioactive dice introduced by Murray and Hart, while preserving its inherent randomness. However, it must be emphasized that the disintegration of radioactive nuclei (radioisotopes) is a completely random phenomenon. The studies mentioned above and the ones without chance involved: They were studied under a deterministic macroscopic model, using the exponential mathematical decay law [4,10,11] at least as a first approximation for… [10,11]. On the other hand, a statistical approximation was carried out by using probability distributions such as binomial or Poisson [12,13]. Other studies use stochastic processes by *master equations* to deduce the respective distributions [7,14,15]. Moreover, there are works related to this innovative methodology, based on the generation of statistical moments and binomial series, that have already been carried out for advanced topics related to multi-fractal analyses and time series (see Refs. [16,17]).

However, it must be emphasized that the disintegration of radioactive nuclei (radioisotopes) is a completely random phenomenon. The studies mentioned above, both deterministic and those incorporating chance, have been conducted. Deterministic macroscopic models have utilized the exponential mathematical decay law [4,10,11], at least as a first approximation for [10,11]. Conversely, statistical approximations have been made using probability distributions such as binomial or Poisson [12,13]. Various studies have employed stochastic processes through *master equations* to deduce the respective distributions [7,14,15]. Furthermore, novel approaches have been devised, based on the generation of statistical moments and binomial series, for advanced topics about multi-fractal analyses and time series (refer to Refs. [16,17]).

The interest in this work lies in conducting a study employing the approach of a stochastic model for radioactive decay. A stochastic model in nuclear physics is a mathematical model that utilizes probability and statistics to describe the behavior of nuclear systems. These models are employed to investigate the dynamics and evolution of nuclear systems, including radioactive decay, fission, and fusion reactions, as well as the behavior of nucleons within the nucleus. They can also be utilized to simulate and predict the behavior of nuclear systems under various conditions and to analyze experimental data. Several commonly employed techniques in stochastic modeling in nuclear physics encompass Monte Carlo simulations, Markov Chain-Monte Carlo, and the Metropolis–Hastings algorithm, as exemplified in [3] part *III* and [18] part *V*.

Based on the above, the purpose of this work is to provide a more realistic approximation to the radioactive decay system by studying its evolution over time using a stochastic approach. In a previous study [9], we generalized the MGF method, as utilized in [8], to encompass scenarios involving the rolling of n dice with s-sides, thereby allowing for real —and even imaginary— numbers. This extension facilitated the representation of a discrete random variable (d.r.v.), which signifies the sum of the resulting values on the top faces. Consequently, we derived the total probability distributions (histograms) that emerge from rolling n multi-sided dice.

Moreover, we employed the MGF method to model population systems of radioactive atoms or nuclei (radioisotopes) undergoing decay in a more realistic context, assuming their behavior to be random —or partially random—. We modeled this phenomenon as a stochastic process using a *Markov (process) chain* and considered it as a *memoryless process*, involving binomial, Poisson, and exponential models. In this paper, we calculated and compared their statistics and how their distributions evolve, simulating with some real physical parameters and periods involving the half-life of the respective radioisotopes used in nuclear medical physics (see, for example [12,13,19,20]).

We have further enriched our understanding by calculating additional statistical quantities, such as moments, intrinsic fluctuations, and *entropy metrics* associated with binomial and Poisson distributions. This analysis, a first in the field, accounts for the random nature of these physical systems and provides deeper insights into the complete system. Specifically, we computed the entropy using *Shannon's information theory* [21], treating the distributions as stochastic processes. This approach, which had not been reported previously, allows us to track the evolution of system configurations by comparing the respective entropy metrics.

Our study proceeds as follows: In Section 2, we briefly introduced to radioactive decay phenomena and toy models from an ambivalent viewpoint, that is, deterministic and non-deterministic, Analyzing the models mentioned previously. Section 3, gives a quick overview of so-called stochastic processes. In the first instance, we showed the previously studied models of the Binomial

Process 3.2; which we needed as a point of comparison for our stochastic study. In Section 4, we carried out a study of the so-called master equations (Kolmogorov) to model two-state processes 4.2, and Poisson 4.3. Then, we modeled the Poisson process from the perspective of the *Birth-Death processes* separately —*Pure Death Process* (**PDP**) 4.3 and *Pure Birth Process* (**PBP**) 4.4— by introducing the concept of *sojourn time* into the radioactive system. It is shown 4.4.1 that partial nuclear decay can occur as the system evolves in different periods and with different parameters. In Section 5, we showed our main results obtained, namely, the relative fluctuations of the random system (using the *Fano Factor* 5.1) and the calculation of the *Stochastic Entropy* 5.2 for the binomial and Poisson processes. Following the common thread of the didactic research [4–7], we developed a deeper approach to radioactive decay, leaving the bases to carry on with this investigation and other physical systems with the same stochastic approach and their entropy for future work. Finally, in Section 6, we gave our conclusion.

2. Radioactive decay

Radioactive decay is, as we mentioned earlier, essentially a random phenomenon in which a nucleus or atom spontaneously emits a particle and becomes a new chemical element. We cannot predict exactly when a given nucleus will suffer this event, but we can study a large collection of nuclei and draw some interesting conclusions related to their statistical and probabilistic properties. Radioactive decay, also known as radioactivity, is a natural phenomenon that occurs in the nucleus of some elements. Radioactive elements are called isotopes each isotope has a specific rate of decay, which means that it can be predicted with some accuracy how many isotopes will decay in a given period. Some isotopes have a very slow rate of decay and can remain stable for thousands of years, others have a very fast rate of decay and disintegrate in a matter of seconds or minutes. When an isotope decays, it emits one of three subatomic particles: an electron, a proton, or a neutron. Decay involving the emission of electrons is called beta decay, while decay involving the emission of protons is called alpha decay.

Decay involving the emission of neutrons is called gamma decay. Each type of decay has a different effect on the atomic nucleus and can create a different element. Radioactive decay is an important process in nature and has several applications in everyday life. For example, it is used in medicine to produce images of the human body and to treat some cancer types. It is also used in industry to detect equipment failures and to produce nuclear energy. However, exposure to high levels of radiation can be dangerous and cause health damage. Therefore, it is important to take precautions when working with radioactive materials (see for example: [10,11,13,20,22]). On the other hand, in this paper, as we already mentioned, we do not perform an analysis of the properties and conditions necessary for the physical detection of radiation or particles involved in radioactive decay, nor experimental statistics. For a study on this matter, those interested can consult: [10,11,13,19,20,22,23] and the additional sources cited therein.

Therefore, when we apply stochastic process modeling to the detection of photons or radioactive isotope radiation, the detection count probabilities will largely depend on the periods, physical parameters, type of detection device of the particles, and so on. However, the statistics associated with particle counting —for example: random polychromatic natural light or photons of monochromatic light (such as a Laser)— depend on the type of radiation, namely classical or non-classical (quantum-bosons), in addition to taking into account whether it is coherent, incoherent or partially coherent light. These characteristics of counting and experimental measurement of radiation can be transferred to the nuclear context taking into account its particularities such as the type of particles (fermions) emitted and absorbed (alpha, beta, etc.). Because in general, the light has a random behavior due to the unpredictable fluctuations of the sources and propagation medium. However, if we consider that random fluctuations of intensity are relatively low, then Poisson's statistics are ideal for this type of process (Optical Coherence Theory; see refs:*Mandel and Wolf* [2] chapters 11 and 12, and the additional sources cited therein).

2.1. Radioactive decay: Toy model

A toy model is a simplified model or representation of a system, concept, or phenomenon that is used to illustrate or explain complex ideas in a more manageable and easily understood way. Toy models are often used in scientific research and education to explore the basic principles of a system or to test hypotheses and predictions. Toy models are usually designed to capture the essential features of a system while ignoring or simplifying other details that are not directly relevant to the topic being studied. This allows researchers to focus on the fundamental mechanisms or processes, rather than becoming entangled in unnecessary complexity. Toy models can be used in a wide range of disciplines, including physics, biology, economics, and computer science. They are often used to study complex systems that are difficult to analyze using traditional mathematical or computational methods. Some examples of toy models include: 1. The typical model, of a simple pendulum, is used to illustrate the basic dynamic principles of small oscillations. 2. Model a simple economic system, which is used to explore the effects of different economic policies or market conditions. 3. A neural network model, used to study the basic principles of learning and decision-making in artificial intelligence systems. 4. A simplified model of radioactive decay, characterized by unbiased dice rolls (although we could include intrinsic bias). This is a considerable physical and mathematical simplification; however, retaining its random traits, brings us closer to the reality of this natural process. Toy models can be very useful for getting a basic insight into a system or concept; however, they should not be confused with real-world systems, which are often much more complex and can exhibit behavior that the toy model does not take into account. Example 4, would be the one that interests us most, as it aligns with the focus of our research on stochastic modeling. In [4] Murray and Hart introduce a toy model for didactically studying the phenomenon of radioactive decay. Based on the rolling of unbiased 6-sided dice, along with the exponential (deterministic) law of decay, they carry out a Taylor series expansion of the binomial (geometric progression) and exponential functions. By comparing the two series, they show how to approximate the exact values that describe the disintegration process of radioactive atoms (nuclei). Essentially, they propose the model (binomial geometric progression):

$$N_n = 1000(1 - 1/6)^n,$$

where N_n is the total number of radioactive dice, with $N_0 = 1000$ is the initial number, which can vary. And n represents the number of mass throws of the radioactive dice. Later, they make the comparison with the decay law:

$$N_t = 1000 \exp(-t/6).$$

Here, N_t is a function (of time), which represents the exponential decay of the radioactive dice.

As we already mentioned, they compare the respective Taylor series of both formulas, reaching a very interesting conclusion (which they no longer develop theoretically): that the way for both series to converge is to make the number of dice faces increase, which they represent with the variable p. This allows for greater precision in the behavior of the real phenomenon. Under this premise, Sánchez et al. [9], generalize the toy model (although it is still an approximation) using the moment-generating function method (MGF-M). In the paper by Sánchez et al. [9], they generalize of the toy model (6-sided dice) to s-sided dice. That is, for an arbitrary number of sides. Considering its random characteristics, as an intrinsic feature in the throwing of dice for games of chance, as used in the article by Singh, et al. [8] for gambling in *Las Vegas casinos*. In the paper by Sanchez, they expand and generalize Singh's method to apply not only to traditional 6-sided dice but also to *multi-sided dice* (s-sided) and the didactic concept of radioactive dice under the idea proposed by Murray et al. Their application is not limited to integer values, but also extends to all real numbers and even imaginary units. They got the model:

As previously mentioned, they compare the respective Taylor series of both formulas, reaching a very interesting conclusion (which they no longer develop theoretically): that the way for both series to converge is to make the number of dice faces increase, which they represent with the variable p. This allows for greater precision in the behavior of the real phenomenon. Under this premise, Sánchez et al. ([9]), use the moment-generating function method (MGF-M), although it is still an approximation. Generalize the toy model (6-sided dice) to s-sided dice. That is, for an arbitrary number of sides. Considering its random characteristics, as an intrinsic feature in the throwing of dice for games of chance, as used in the article by Singh, et al. ([8]) for gambling in Las Vegas casinos. In the paper by Sanchez, they expand and generalize Singh's method to apply not only to traditional 6-sided dice but also to multi-sided dice (s-sided) and the didactic concept of radioactive dice under the idea proposed by Murray et al. Their application is not limited to integer values but also extends to all real numbers and even imaginary units. They described the model as:

$$M_{\Sigma}(t^*) = \frac{1}{j^n}\left(\sum_{r_j \in \mathbb{R}}^{S_{j-1}} e^{kt^*}\right)^n = \left(\sum_{r_j \in \mathbb{R}}^{S_{j-1}} \frac{1}{j} e^{r_j t^*}\right)^n.$$

where $M_{\Sigma}(t^*)$ is the moment-generating function (MGF), j is an integer ($j \in \mathbb{Z}^+$) that labels the number of sides (discrete random variable **d.r.v.**) of the *multi-sided die*, and S_{j-1} and $r_j \in \mathbb{R}$, (t^* does not represent time, it is an auxiliary or dummy parameter used in elementary probability theory). In addition, gives the probabilistic weight of the number of sides (i.e., $p_j = p(j) = 1/j$). r_j can range from negative values, zero and for any real number and imaginary. The formula they obtain, in addition to being very general, encompasses the probabilistic and statistical features of the random nature of the dice system.

According to this generalization, we can approximate the radioactive toy model based on 6-sided dice to a stochastic process modeled by probability distributions, fundamentally binomial, and Poisson. Nevertheless, we could choose another model with stochastic characteristics to study the phenomenon of radioactive decay. For example, using stochastic differential equations, whereby our random variables and time are both continuous, including so-called Gaussian noise. But the essential idea of our study is to analyze the behavior and performance of the binomial and Poisson distributions, as stochastic models for radioactive decay.

2.2. Radioactive decay: deterministic model

Deterministic models are those in which the outcome is completely determined by the initial conditions and rules of the model. On the other hand, random models include elements of chance or uncertainty, meaning that the outcome is not completely determined by the initial conditions and rules of the model. For example, imagine a model that simulates the parabolic motion of a particle using Newtonian laws. A deterministic model of this system could accurately predict the position of the particle at any given time, as long as the initial conditions (such as the initial position, velocity, and angle of the particle) and the rules of the model (such as the law of gravity) are known. However, if the model includes random elements, such as an unpredictable gust of wind acting on the particle, then the final result would not be completely determined by the initial conditions and rules of the model.

In general, deterministic models are easier to understand and predict. Still, they may be less accurate than random models in situations where there is uncertainty or unknown randomness, such as *white noise* (see for example [18] chap. 17 and [2] chap. 2). Random models, on the other hand, can be more accurate in these situations. Still, they may also be more difficult to understand and predict due to the uncertainty or chance involved. In the context of radioactive decay, we know that it is a completely random phenomenon. Since the disintegration of nuclei is completely spontaneous, its behavior corresponds to quantum and nuclear laws.

However, it has also been studied in the context of a deterministic macroscopic model using the exponential decay law (see Eq. (1)), to make approximations and simplifications of these types of systems. [10,11,20]. The exponential decay law is a mathematical relationship that describes how a quantity decreases exponentially over time. This relationship can be expressed by Eq. (1), where N_0 is the initial amount of the substance, $N(t)$ is the final amount of the substance, t is the elapsed time, and γ is an (inverse) time constant characteristic of the material type. In the context of a deterministic macroscopic model, this law is used to describe how systems change over time in a predictable and deterministic way. For example, it can be used to describe how a concentration of a substance dissipates in the environment, or how a radioactive material disintegrates. The study of radioactive decay continues to be a topic of current interest. For example, Recently, A. Malyzhenkov, V. Lebedev, and Alonso Castro [24] carried out an interesting research on what they call the *Nuclear Decay Recoil (and nuclear recoil spectroscopy)*, making use of the exponential decay model, and some innovative experimental techniques related with optical traps to capture and levitate particles during radioactive kickback.

We must bear in mind that according to all this, we have two approaches, namely, deterministic and random, to describe the phenomenon of radioactive decay. Each with its respective advantages and disadvantages, without losing sight of the fact that the phenomenon is one hundred percent random. Therefore, it is reasonable to consider the stochastic model to be a more accurate approximation of reality. The second one is given through a random variable associated with its probability distribution function, but the exponential model must be given the appropriate interpretation, as mentioned by S. P. Huestis [7]. Nevertheless, it is not the only interpretation that supports the random process of radioactive decay, there are also opinions without including the deterministic one, where studies (see, [12,14,15]) are made directly on the different types of distributions that describe the process, particularly binomial and Poisson distributions.

Mathematically deterministic description: The process when an isotope A transforms to isotope B, written as $I_A \longmapsto I_B$ is represented by an ordinary linear differential equation given by

$$\frac{dN_{I_A}(t)}{dt} = -\gamma N_{I_A}(t);$$

where the solution is ($N_{I_A}(0) = N_0$)

$$N_{I_A} \equiv N(t) = N_0 e^{-\gamma t} \tag{1}$$

where $N(t)$ is interpreted as the number of undecayed nuclei remaining after a time t, N_0 is the number of nuclei present at time $t = 0$ and γ is the decay constant specific of each isotope given by $\gamma = 1/\tau$, where $t_{1/2} = \ln 2/\Gamma = (\ln 2)\tau$ are the mean lifetime and the half-life of the decaying atoms, respectively [10,11]. The Eq. (1) can be interpreted from two different but complementary points of view. One deterministic, in which $N(t)$ represents the total number of isotopes that have not yet decayed, and another interpretation where the expected or average values were considered according the Ref. [7]. The nuclei disintegration process is one by one and the time in which I_A transforms into I_B is random in nature, therefore, the most appropriate way to model it is through a stochastic model.

We should mention that the validity of the Radioactive Decay Exponential Law has been debated for a long time due to notable deviations, not only of a theoretical nature but also when very precise measurements have been conducted. Previous studies [25,26] have addressed this issue. Additionally, experiments by E. B. Norman, et al. have been conducted over extremely precise periods ranging from 0.01% of a half-life to periods encompassing 45 half-lives. These experiments found no significant deviations that would challenge the exponential decay law. Quoting E. B. Norman, et al.: *Our data were also analyzed to search for non-exponential effects of this type. No indications of such deviations were found in our data and the limits derived on their amplitudes are shown in Table II. (see Ref. [26] page 2248)*.

It's worth noting that there has been significant controversy regarding which type of probability distribution is most suitable for a comprehensive statistical description of the radioactive process, particularly in fields such as radio chemistry, nuclear engineering, medical physics, and radiological sciences in general [12,14,15,23]. This arises from the varying decay time scales inherent in the decay constant for different types of radioisotopes and times [12,13]. From a probabilistic and statistical perspective, certain probability distributions, such as the binomial and Poisson distributions, are theoretically understood to be mathematically equivalent. In the upcoming sections, our focus will be on investigating the stochastic model for radioactive decay, particularly emphasizing the binomial and Poisson distributions to compare their probabilistic and statistical characteristics. We will explore their evolution, moments, *Fano factor*, and, notably, their *entropy*.

3. Stochastic processes in radioactive decay

Our goal in this and subsequent sections is to demonstrate that the decay of radioactive isotopes, as described by the *deterministic* Eq. (1), is approximately related to the stochastic process of radioactive decay. This stochastic process is described through a probability distribution function (*PDF*) that accurately captures its inherently random nature. They are distributed through a binomial process, akin to rolling radioactive dice with a few nuclei. However, for a more accurate depiction of reality, we adopt a Poisson Process Model (**PPM**). This approach is particularly suitable for situations where the number of trials is very large and the probability per event is small. Nevertheless, it's crucial to note that both the periods of the process and the parameters are also highly relevant.

Thus, we compare it with the Binomial Process Model (**BPM**) as proposed in the Refs. [7,12,14,15]. Theoretically, the Poisson distribution provides a better approximation for large values of n and small probabilities, such as when the number of isotopes (nuclei) increases significantly. This increase can be likened to the number of sides, akin to the number of dice—multi-sided—involved in the theoretical study of the physical event of radioactive decay. Building on the didactic application of radioactive dice, as demonstrated by Murray and Hart [4], along with other didactic works in physics and chemistry [5–7], we can bridge the gap to more fundamental research on the radioactive decay event. This allows us to demonstrate the effectiveness of applied stochastic processes compared to deterministic models. Our initial approach is based on the formulation of multi-sided dice developed in the reference by Sanchez et al. [9] (see sec. 2.1), where the probability distribution functions (in the discrete case) are theoretically derived. In this elementary study, the authors progressively increase both the number of (radioactive) dice and the number of sides per die — mimicking the behavior of radioactive nuclei, as proposed in the study by Murray and Hart [4] 2.1. Consequently, the respective Poisson (Binomial) style distributions can be generated for each roll of radioactive dice (isotopes).

On the other hand, supported by a stochastic process (Poisson or Binomial), which for a very large n tends to behave like a normal distribution according to the *Central Limit Theorem (**CLM**)* [18,27,28]. Therefore, as the number of radioactive dice increases, along with their sides, the distribution tends to become continuous. Due to random events with this type of behavior, its probability distribution function (PDF) begins to describe a behavior increasingly similar to that of a continuous variable. Therefore, the comparison of the binomial and Poisson distributions in elementary probability theory (see, for example, [18,27,28]), as well as in a more advanced context in the realm of stochastic processes, is more synergistic than op-positional. This is because they are complementary distribution functions, as demonstrated in elementary literature. We further illustrate this synergy in this paper by applying them to the aforementioned stochastic model, with particular emphasis on the Poisson model.

A *Poisson Process* is essentially a stochastic process featuring a discrete random variable evolving (whether discrete or continuous). It describes a sequence of arrivals along the real axis and is commonly employed to model phenomena such as the arrival times of events in a system. This process entails randomly distributed appearances over time, space, or some other one-dimensional variable, with the number of occurrences in any non-overlapping interval being statistically independent. The Poisson process is characterized by its monotone non-decreasing function, making it a *counting process*.

This relates to the repetition of an event along a one-dimensional axis, often represented by the time variable. These events manifest as points randomly distributed along the time axis, as previously mentioned. Further development within a broader theoretical context for the Poisson process can be found in Ref. [29]. For a more explicit and didactic treatment, see also Refs. [18,30–32].

It's worth noting that the Poisson process is a specific case of the so-called Birth and Death process [18], which finds extensive utility and application in the physics of many-particle systems and species biology. We will observe and demonstrate that the radioactive decay process can be regarded as a pure death stochastic process. However, it's essential to mention that the Poisson process can also be derived as a pure birth process. In the following sections, we will provide the conceptual and mathematical arguments to obtain the process from both perspectives.

3.1. Stochastic modeling

Stochastic Markov processes in continuous time offer a modeling framework facilitated by the so-called master equations — a term commonly used in statistical physics [2,3], stemming from the Chapman–Kolmogorov equations (**CKE**) involving discrete random variables in both discrete and continuous time (see, for example, [18,31,32]). This section presents the general model for a continuous-time two-state process, which is adequate for modeling both binomial and Poisson processes.

As mentioned earlier, we demonstrate the validity of a general two-state model from first principles, enabling us to investigate radioactive decay through a stochastic process. Before delving into our stochastic Poisson model, we introduce examples of binomial models proposed by Huestis [7] and Foster et al. [14,15]. These binomial and Poisson processes are among the most widely used in the literature on stochastic systems, though they are not the only ones.

3.2. Binomial process model

As mentioned earlier, we now present two stochastic models corresponding to the binomial distribution. Radioactive decay fundamentally operates within the realms of probability and statistics. The likelihood of one or more nuclei disintegrating is governed by probabilistic events. For a single nucleus, the probability (as a function of time) is modeled according to Huestis as [7]: $\frac{dP_{ip}(t)}{dt} = -\gamma P_{ip}(t)$, Where ip represents the rate of change of the *initial process* (in this case, without decay). This is a simple first-order differential equation, with the initial condition $P_{ip}(t = 0) = 1$, reflecting the physical conditions of radioactive decay, where initially there are no decay events at $t = 0$. The solution to this equation is $P_{ip}(t) = e^{-\gamma t}$. Mathematically, this equation and Eq. (1) are equivalent. However, they carry different meanings: the first equation represents the probability that a nucleus has not decayed, while Eq. (1) represents the deterministic event of the number of nuclei that have not decayed.

That is, with the help of the binomial distribution $P_n(x) = \binom{n}{x} p^x q^{n-x}$ we can substitute the solution function $P_{ip}(t) = e^{-\gamma t}$, and we get

$$P_{N_0}(t) = \binom{N_0}{x} P^x(t)\big(1 - P(t)\big)^{N_0-x} = \binom{N_0}{x} e^{-x\gamma t}\big(1 - e^{-\gamma t}\big)^{N_0-x}. \tag{2}$$

where $P_{N_0}(t)$ is a binomial process, and r.v. $n = \mathbf{x}$ is the number of undecayed nuclei I_A at time t. Whose moments in this case are given by $\mu(t) = \mathbf{E}[X(t)] = N_0 e^{-\gamma t}$ and $Var(t) = \sigma(t) = \sqrt{N_0 e^{-\gamma t}(1 - e^{-\gamma t})}$.

However, Foster et al. [14,15] propose an apparently better stochastic approximation, using a (very simplified) master equation, leading to a binomial process. Nevertheless, we must clarify that this representation is characteristic of a so-called *Yule stochastic process*, which is a *pure birth process*, (the general process refers to the *Birth and Death Process (BDP)* [18,30,32,33], which we will see briefly later in the next section). The Yule (birth) process is a model for population growth used in statistical physics and biology. Thus, they propose the model:

$$\frac{dP(x,t)}{dt} = \gamma \left[P(x-1,t)(N_0 - x + 1) - P(x,t)(N_0 - x) \right] \tag{3}$$

With initial conditions $P(0,0) = 1$ and $P(x,0) = 0$ (which are quite arbitrary in this work by Foster et al.). According to Foster et al. [14,15], $P(t)$ must satisfy the fundamental differential equation $P'(t) + \gamma P(t) = \gamma$, with solution $P(t) = (1 - e^{-\gamma t})$, where γ is the radioactive decay parameter (decay rate per unit time). With the following statistical moments: The expectation and variance for $X(t)$ are: $\mathbf{E}(X) = N_0(1 - e^{-\gamma t})$, and $\mathbf{Var}(X) = N_0 e^{-\gamma t}(1 - e^{-\gamma t})$; N_0 is the number of initial radioactive nuclei. Eq. (2) changes apparently little in relation to the exponents of its factors. However, this is substantially deeper and has to do with what we said above about the birth and death process (BDP). Which we explain later in section 4.2. Then the solution they offer (Foster et al.) for the master Eq. (3), obtaining the binomial distribution (stochastic process) is

$$P_{N_0}(t) = \binom{N_0}{x} P^x(t)\left(1 - P(t)\right)^{N_0 - x} = \binom{N_0}{x}\left(1 - e^{-\gamma t}\right)^x e^{-(N_0 - x)\gamma t}. \tag{4}$$

To better insight into the above equations (models) (2) to (4), let us point out the following concepts: Despite the fact that the Huestis model is based on more deterministic and elementary principles, we must say that their model is more accurate and suitable for radioactive decay than the model by Foster et al. Because Huestis's is similar to the model based on a pure stochastic death process —In particular, it is the so-called **linear death process**, where the parameter is proportional to the transition (decay) rate, i.e., $\mu_n = n\mu$.[1] As we demonstrate in Section 4— that it is more suitable for the decay of nuclei. And Foster's et al. despite being based on the theory of master equations, his model is based on a birth process — to be precise, it is based on the so-called *Yule Process*, which describes the growth of a physical or biological population—. This implies that populations (in this case of radioactive nuclei) grow, and do not decrease as happens in the reality. In figures 2-(a) and (b), we show the plots (PDFs) of these two models together with the stochastic models that we propose for the Poisson Process (see Section 4 *y* 4.3). For a fuller discussion of these stochastic processes, see Refs. [3,18,30,32,33].

4. Master equation and stochastic models

In this section, we propose a model based on the Chapman–Kolmogorov equations (**CKE**) for a discrete random variable (**d.r.v.**) in continuous time, known as the *master equations* [2,3,18] in statistical and quantum physics. We do so, specifically, for a two-state or level system, which is enough to get the two stochastic models based on the binomial and Poisson distributions.

4.1. Master equation and Chapman–Kolmogorov equations (CKE)

Chapman–Kolmogorov equations (CKE): The Chapman–Kolmogorov equations are a set of equations that describe the time evolution of a Markov chain. These are stochastic model that describes the random evolution of a system through a series of discrete states (either discrete or continuous time). Suppose that the Markov chain has a set of states $1, 2, \ldots, N$ and that the transition probability of going from state i to state j in one-time step is $p_{i,j}$. Then, the discrete-time Chapman–Kolmogorov equations are expressed as follows:

$$P_{i,j}^{(n+m)} = \sum_{k=1}^{N} P_{i,k}^{(n)} P_{k,j}^{(m)}$$

where $P_{i,j}^{(n)}$ is the probability that the chain will be in state j after n time steps, starting from state i. The Chapman–Kolmogorov equations are based on the idea that the probability of reaching a final state after $n+m$ steps can be resolved into the sum of the probabilities of reaching each possible intermediate state after n steps, and then reach the final state after m steps. This equation is important for the analysis of Markov chains, as it allows us to calculate the probability that the chain will be in any state after an arbitrary number of steps, given knowledge of the one-step transition probabilities.

Now we define a *Transition probability matrix function* (**TPMF**) for a Continuous-time Markov chain (**CTMC**) $X(t)$ (see [18] chap.16, [32] chap.6, [33] chap.VI, and [30]) chap.3 with the conditional probabilities

$$P_{ij}(t) = \mathbf{P}[X(s+t) = j | X(s) = i]; \quad i, j \in \mathcal{S}, \quad 0 \le t < \infty$$

where $\mathcal{S}$ is the state space. $P_{ij}(t)$ are called transition probability functions, and $\mathbf{P} = [P_{ij}(t)]$, is called the transition probability matrix function (**TPMF**). The Chapman–Kolmogorov discrete (**d.r.v.**) equations take the continuous-time form:

$$P_{ik}(s+t) = \sum_{j \in \mathcal{S}} P_{ij}(s) P_{jk}(t), \quad i, k \in \mathcal{S}, \quad s, t \ge 0 \tag{5}$$

And in matrix form they are: $\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t)$. Making the change in Eq. (5) of s by an infinitesimal interval h, we obtain: $P_{ik}(h+t) = \sum_{j \neq k} P_{ij}(h) P_{jk}(t) + P_{ii}(h) P_{ik}(t)$. So by carrying out the limit process, when $h \to \infty$, we get for the case of a continuous-time

[1] The γ, λ, and μ parameters can be used similarly in this reference Section 3.2 (as well as 4.4.2). However, starting from Section 4.2, we make a distinction, especially between the λ_i and μ_i parameters for the different *Birth and Death processes*.

stochastic process or **CTMC**, the Chapman–Kolmogorov differential equations: Referred to in the specialized stochastic literature as the **Kolmogorov's backward differential equation** (**KBDE**)[2]

$$\begin{aligned}\frac{d}{dt}P_{ij}(t) &= \sum_{k\neq j} q_{ik}P_{kj}(t)+q_{ji}P_{ij}(t)=\sum_{k\in S} q_{ik}P_{kj}(t), \quad i,j\in S, \quad 0\le t<\infty \\ \frac{d\mathbf{P}(t)}{dt} &= \mathbf{QP} \quad \text{in matrix form}\end{aligned} \tag{6}$$

where $P_{ij}(t)$ is the probability that the stochastic process passes from state i to state j in time t, and q_{ik} is the state transition rate i to state k. The first addition on the right-hand side of Eq. (6) represents the probability that the stochastic process will go from state i to any other state k and then to state j, while the second addition represents the probability that the stochastic process stays in state j and then goes to state i. The solution methods for solving Kolmogorov equations are varied. We focus on solving them using matrix techniques [18,34]. However, in the references cited in this paper [3,18,30,32,33], you can find the different standard solution methods for these systems. A very general solution to Eq. (6) is based on the exponential of a matrix [18,34], as follows

$$\mathbf{P}(t)=\exp(\mathbf{Q}t)=e^{\mathbf{Q}t}=\sum_{n=0}^{\infty}\frac{[\mathbf{Q}t]^n}{n!}=\mathbf{I}+\mathbf{Q}t+\frac{\mathbf{Q}^2t^2}{2!}+\frac{\mathbf{Q}^3t^3}{3!}+\cdots \quad t\ge 0 \tag{7}$$

where $\mathbf{P}(0)=\mathbf{I}$ is the identity matrix, or $P_{ij}(0)=\delta_{ij}$ the diagonal elements of the matrix. In addition, it is required that $\mathbf{Q}$ be a uniform matrix in order for it to converge. In the next section, we develop the two-state model (*two-state Markov Chain*) and its general solution.

4.2. Master equation: Two-state Markov process (TSMP)

Here we present the model of a two-state Markov Chain Process $\{\mathbf{X}(t)\}$, using the master equation (Kolmogorov equation). We obtain its general solution (in Appendix B) and generalize it to N independent and indistinguishable two-state Markov Processes, which gives us the master equation (model) for the binomial process and also for the so-called Birth-Death Process (**BDP**), which the Poisson process can be derived.

TSMP System as Birth-Death Process (BDP): We can consider that a process of births and deaths is represented by the difference of two states, i.e., $\mathbf{N}(t)=\mathbf{S_a}(t)-\mathbf{S_d}(t)$, where $\mathbf{S_a}(t)$ corresponds to an open switch (arrival counting process or open channel). and $\mathbf{S_d}(t)$ to a switch closure (departure counting process or channel closed or busy). This system is governed by the following system of stochastic equations (*Kolmogorov backward equations*), represented with the help of (6). Which can be represented in matrix form with the corresponding transition rates (see (9)), namely (see Refs. [3,18,30,32,33]),

$$\begin{aligned}\frac{dP_{0j}(t)}{dt} &= -\lambda_0 P_{0j}(t)+\lambda_0 P_{1j}(t) \\ \frac{dP_{ij}(t)}{dt} &= \mu_i P_{i-1,j}(t)-\left(\lambda_i+\mu_i\right)P_{i,j}(t)+\lambda_i P_{i+1,j}(t), \quad i\ge 1,\end{aligned} \tag{8}$$

where the boundary conditions are $P_{i,j}(0)=\delta_{ij}$. And the parameters λ_i and μ_i are called, respectively, the *infinitesimal birth and death rates*.

We consider the two-state Markov Chain $\{\mathbf{X}(t)\}$ with states: $S=\{\mathbf{s}_0,\mathbf{s}_1\}$, where state $\mathbf{s}_0$ represents that the channel is active and state $\mathbf{s}_1$ represents that the channel is inactive or busy. Both the duration of the period of the active and busy states are random, independent, exponential variables (they follow an exponential distribution) with α and β parameters respectively. Corresponding to the transitions, $\mathbf{Q}(\mathbf{s}_0\to\mathbf{s}_1)=\alpha$ and $\mathbf{Q}(\mathbf{s}_1\to\mathbf{s}_0)=\beta$. To solve this system, we need to establish the transition probability matrix function (**TPMF**) $\mathbf{P}=[P_{ij}(t)]$, using its infinitesimal generator $\mathbf{Q}$[34]

The probability of transitions in a time interval Δt occur as follows: the probability (array elements) $p_{ij}(t)$ of de system going from $\mathbf{s}_0$ to $\mathbf{s}_1$ in interval Δt is $p_{01}(\Delta t)=\lambda\Delta t+o(t)$, and similary for $p_{10}(\Delta t)=\mu\Delta t+o(t)$. From Eq. (8), with $\lambda_0=\alpha$, $\lambda_i=0$, $\mu_i=0$, for $i\neq 0$ and $\mu_1=\beta$. We obtain the differential equation system

$$\begin{aligned}\frac{dp_{00}(t)}{dt} &= -\alpha p_{00}(t)+\alpha p_{10}(t) \\ \frac{dp_{10}(t)}{dt} &= \beta p_{00}(t)-\beta p_{10}(t),\end{aligned}$$

This system have the infinitesimal generator matrix of the **CTMF** expressed by

$$\mathbf{Q}=\begin{pmatrix}-\alpha & \alpha \\ \beta & -\beta\end{pmatrix} \tag{9}$$

[2] Kolmogorov's forward differential equation (**KFDE**) also exists, which in matrix form is given by $\frac{d\mathbf{P}(t)}{dt}=\mathbf{PQ}$. This is derived similarly to Kolmogorov's backward differential equation (**KBDE**) (6), (see Refs. [18,30,32,33]). However, the solutions are practically similar, so in Sections 4.1 and 4.2, we will only use Eq. (6).

[3] The infinitesimal generator matrix is defined by $\mathbf{Q}=\frac{d\mathbf{P}(0)}{dt}$; and we can define it simply as the matrix $\mathbf{Q}=[Q_{ij}(t)]$, (see [18] chap.16), and [3] chap.3.

[4] Remember that in elementary stochastic modeling, a two-state Markov Chain is represented by the matrix $\mathbf{P}=\begin{pmatrix}1-\alpha & \alpha \\ \beta & 1-\beta\end{pmatrix}$, which is a stationary particular case of a $\mathbf{P}(t)$ distribution for two states, (not to be confused with the matrix $\mathbf{Q}$).

The **TPMF**, $\mathbf{P} = [P_{ij}(t)]$ for small times Δt can be expressed by

$$\mathbf{P}(t) = e^{\mathbf{Q}\Delta t} = \mathbf{I} + \mathbf{Q}\Delta t + o(t)\mathbf{E} = \begin{bmatrix} 1-\alpha\Delta t + o(t) & \alpha\Delta t + o(t) \\ \beta\Delta t + o(t) & 1-\beta\Delta t + o(t) \end{bmatrix} \tag{10}$$

where $\mathbf{E}$ is square matrix of all ones. And $o(t)$ is defined as a quantity that tends to zero faster than t, as $t \to 0$; i.e., $\lim_{t\to 0}[o(t)/t] = 0$. To solve this system we can use the *spectral expansion technique* [18,34], for the infinitesimal generator $\mathbf{Q}$. Which we can expand the infinitesimal generator as $\mathbf{Q} = \mathbf{U\Lambda U^{-1}} = \mathbf{U\Lambda V} = \sum_{j\in S} \lambda_j \mathbf{E}_j$; where $\mathbf{\Lambda}$ is a diagonal matrix represented by $\mathbf{\Lambda} = diag[\lambda_0, \lambda_1, \lambda_2, \ldots]$, and λ_j are the *eingenvalues* of matrix $\mathbf{Q}$; i.e., $det|\mathbf{Q} - \lambda_j \mathbf{I}| = 0$, and $\mathbf{E}_j = \mathbf{u}_j \mathbf{v}_j^T$, $(\mathbf{UV}^{-T} = \mathbf{I})$ are the projection matrices, with $j \in S$. the similarity matrix $\mathbf{U}$, contains the right-eigenvector column $\mathbf{u}_j$, generated by the eigenvalue λ_j. Likewise, the jth row vector $\mathbf{v}_j$ of $\mathbf{V} = \mathbf{U}^{-1}$ corresponds to the left-eigenvector of the eigenvalue λ_j.

Solution by spectral expansion: The equation for the evolution of the probability transition vector is

$$\frac{d\mathbf{p}(t)}{dt} = \mathbf{Qp}, \quad \text{where} \quad \mathbf{p}(t) = \begin{bmatrix} p_0(t) \\ p_1(t) \end{bmatrix}$$

And the matrix $\mathbf{Q}$ is defined in (9). With formal general solution

$$\mathbf{p}(t) = e^{\mathbf{Q}t}\mathbf{p}(0), \quad \text{where} \quad \mathbf{p}(0) \quad \text{are the initial conditions.}$$

It can be shown [18,34] that the matrix $\mathbf{Q}$ and $\mathbf{\Lambda}$ admit a diagonal exponential solution of the following form

$$\mathbf{P}(t) = e^{\mathbf{Q}t} = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{Q}^n = \sum_{n=0}^{\infty} \frac{t^n}{n!}\left[\mathbf{U\Lambda U^{-1}}\right]^n \quad \Longrightarrow \quad e^{\mathbf{Q}t} = \mathbf{U}\sum_{n=0}^{\infty} \frac{t^n}{n!}\left[\mathbf{\Lambda}\right]^n \mathbf{U}^{-1} = \mathbf{U}e^{\mathbf{\Lambda}t}\mathbf{U}^{-1} \tag{11}$$

Also $\mathbf{U}$ is a similarity matrix. To solve, we form the system of differential equations composed of the matrix $\mathbf{Q}$ and the column probability vector $\mathbf{p}(t) = (p_0(t), p_1(t))^T$, with the general initial condition: $\mathbf{p}(0) = (p, 1-p)^T$; and normalization condition $p_0(t)+p_1(t) = 1$. In the particular case of the infinitesimal generator $\mathbf{Q}$ of our two-state system, we obtain the following

$$det|\mathbf{Q} - \alpha_j \mathbf{I}| = det \begin{bmatrix} -\alpha-\lambda & \alpha \\ \beta & -\beta-\lambda \end{bmatrix} = 0 \tag{12}$$

We get $\lambda(\lambda+\alpha+\beta) = \lambda^2 + \lambda\beta + \lambda\alpha = 0$. From this, we can quickly verify that the eigenvalues of $\mathbf{Q}$ are $\lambda_0 = 0$ and $\lambda_1 = -(\alpha+\beta)$. With the respective eigenvectors given by

$$\mathbf{u}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \qquad \mathbf{u}_1 = \begin{bmatrix} \alpha \\ -\beta \end{bmatrix} \tag{13}$$

The matrix $\mathbf{U}$ composed of the eigenvectors and its inverse $\mathbf{U}^{-1}$ make up the matrix $\mathbf{Q}$, as follows

$$\mathbf{Q} = \mathbf{U\Lambda U^{-1}} = \begin{pmatrix} 1 & \alpha \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -(\alpha+\beta) \end{pmatrix} \frac{1}{(\alpha+\beta)} \begin{pmatrix} \beta & \alpha \\ 1 & -1 \end{pmatrix} \tag{14}$$

Obtaining the general exponential solution (**TPMF**) of Eq. (10)

$$\begin{aligned}
\mathbf{P}(t) &= \mathbf{U}\begin{pmatrix} 1 & 0 \\ 0 & e^{-(\alpha+\beta)t} \end{pmatrix}\mathbf{U}^{-1} = \frac{1}{(\alpha+\beta)}\begin{bmatrix} \beta+\alpha e^{-(\alpha+\beta)t} & \alpha-\alpha e^{-(\alpha+\beta)t} \\ \beta-\beta e^{-(\alpha+\beta)t} & \alpha+\beta e^{-(\alpha+\beta)t} \end{bmatrix} \\
&= \frac{1}{(\alpha+\beta)}\begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix} \\
\mathbf{P}(t) &= \frac{1}{(\alpha+\beta)}\begin{bmatrix} \alpha+\beta & 0 \\ 0 & \alpha+\beta \end{bmatrix} + \frac{1}{(\alpha+\beta)}\begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix} \\
\mathbf{P}(t) &= \mathbf{I} + \frac{1}{(\alpha+\beta)}\mathbf{Q} - \frac{e^{-(\alpha+\beta)t}}{(\alpha+\beta)}\mathbf{Q} = \mathbf{I} + \frac{1}{(\alpha+\beta)}\left(1 - e^{-(\alpha+\beta)t}\right)\mathbf{Q}
\end{aligned} \tag{15}$$

where the solutions for initial conditions: $p_0(0) = p_{00}(0) = 1$, and $p_1(0) = p_{10}(0) = 0$, corresponding for states $S = \{\mathbf{s}_0, \mathbf{s}_1\}$, Remember that they refer to an active state — the first one — and an inactive state — the second one —. Nevertheless, the opposite initial conditions can be chosen, i.e. $p_0(0) = p_{01}(0) = 0$, and $p_1(0) = p_{11}(0) = 1$, which would be given by the second column of the solution matrix (15)

$$\begin{aligned}
p_0(t) &= p_{00}(t) = \frac{1}{(\alpha+\beta)}\left[\beta+\alpha e^{-(\alpha+\beta)t}\right] = \frac{\beta}{(\alpha+\beta)} + \frac{\alpha}{(\alpha+\beta)}e^{-(\alpha+\beta)t} \\
p_1(t) &= p_{10}(t) = \frac{1}{(\alpha+\beta)}\left[\beta-\beta e^{-(\alpha+\beta)t}\right] = \frac{\beta}{(\alpha+\beta)} - \frac{\beta}{(\alpha+\beta)}e^{-(\alpha+\beta)t}
\end{aligned} \tag{16}$$

Eqs. (16) are complemented by those mentioned in the text (obtained in the **TPMF** matrix (14)), that is,

$$\begin{aligned}
p_{01}(t) &= \frac{1}{(\alpha+\beta)}\left[\alpha-\alpha e^{-(\alpha+\beta)t}\right] = \frac{\alpha}{(\alpha+\beta)} - \frac{\alpha}{(\alpha+\beta)}e^{-(\alpha+\beta)t} \\
p_{11}(t) &= \frac{1}{(\alpha+\beta)}\left[\alpha+\beta e^{-(\alpha+\beta)t}\right] = \frac{\alpha}{(\alpha+\beta)} + \frac{\beta}{(\alpha+\beta)}e^{-(\alpha+\beta)t}
\end{aligned} \tag{17}$$

To find the *stationary distribution* —that is, *steady and invariant*—, which allows us to understand the behavior in the long term, we apply the limit,

$$\lim_{t\to\infty} P_{ij}(t) = \pi(j) = \pi_j$$

and parameterize as follows (with the help of the two states $\{\mathbf{s}_0 = 0, \mathbf{s}_1 = 1\}$): $\pi_0 = \pi(0) = \alpha/(\alpha+\beta)$, $\pi_1 =$
$= \pi(1) = \beta/(\alpha+\beta)$ and $\theta = \alpha+\beta$, which represent the *stationary distributions*. So, by (16) we get

$$\begin{aligned} p_{00}(t) &= \pi(1) + \pi(0)e^{-(\alpha+\beta)t} = \left(1-\pi_0\right) + \pi_0 e^{-\theta t} \\ p_{10}(t) &= (1-\pi(0)) - \left(1-\pi(0)\right)e^{-(\alpha+\beta)t} = \left(1-\pi_0\right)(1-e^{-\theta t}) \end{aligned} \tag{18}$$

Similarity

$$\begin{aligned} p_{01}(t) &= \pi(0) + \pi(0)e^{-(\alpha+\beta)t} = \pi_0(1-e^{-\theta t}) \\ p_{11}(t) &= \pi(0) + \left(1-\pi(0)\right)e^{-(\alpha+\beta)t} = \pi_0 + \left(1-\pi_0\right)e^{-\theta t} \end{aligned} \tag{19}$$

We can observe, that

$$\lim_{t\to\infty} p_{01}(t) = \lim_{t\to\infty} p_{11}(t) = \pi(0) = \pi_0$$

So we have found the *long-run probability* is π_0, i.e., the stationary probability, which is independent of where the process starts. Also, its probabilistic complements:

$$\lim_{t\to\infty} p_{00}(t) = \lim_{t\to\infty} p_{10}(t) = \pi(0) = 1-\pi_0$$

where we use the normalization condition $\pi_0 + \pi_1 = 1$, (Kolmogorov's first axiom: $\sum_{j=1}^{\infty} \pi_j = 1$).

*4.3. Master equation: Poisson's process as a pure death process **(PDP)***

Here we introduce the Kolmogorov's forward differential equation (**KFDE**) in its general form — similarly to (**KBDE**) (8) — and for the particular case of the Poisson process — as a *Pure Death Process* (**PDP**)— for our radioactive decay system. However, this is a fairly general formulation applicable to other systems of the Poisson type.

$$\begin{aligned} \frac{dP_{0j}(t)}{dt} &= -\lambda_0 P_{0j}(t) + \mu_1 P_{1j}(t) \\ \frac{dP_{ij}(t)}{dt} &= \lambda_{j-1}P_{i,j-1}(t) - \left(\lambda_j + \mu_j\right)P_{i,j}(t) + \mu_{j+1}P_{i,j+1}(t), \quad j \geq 1, \end{aligned} \tag{20}$$

Pure Death Process (**PDP**): In fact, both in the **KFDE** (20) and (**KBDE**) (8), we have that the general equations are the second, from which we can extract all the information of the system. For a **PDP**-type Poisson process, we have: from the second equation of (20) we make the following changes, —including the boundary conditions at $n = 0$ and $n = N_0$ —: $\lambda_j = \lambda_{j-1} = 0$ and $\mu_j = \mu_{j+1} = \mu_{n+1} = \mu$. Also $P_{i,j}(t) = P_n(t)$, and $P_{i,j+1}(t) = P_{n+1}(t)$. We have that parameter (*transition rate*) $\mu = 1/\tau$, i.e., decay rate per unit time.

$$\begin{aligned} \frac{dP_{N_0}(t)}{dt} &= -\mu P_{N_0}(t) = -\frac{1}{\tau}P_{N_0}(t) \\ \frac{dP_n(t)}{dt} &= -\mu P_n(t) + \mu P_{n+1}(t), \quad N_0 - 1 \geq n > 1, \\ \frac{dP_0(t)}{dt} &= \mu P_1(t) = \frac{1}{\tau}P_1(t) \end{aligned} \tag{21}$$

the initial probability distribution is $\pi(0) = P_{n=N_0}(t=0) = \delta_{N,N_0} = \delta_{n,n_0}$. Solving the first equation of (21), we obtain: $P_{N_0}(t) = e^{-\mu t} = e^{-t/\tau}$. So, the time-dependent solution for $P_n(t)$ is given by

$$\begin{aligned} P_n(t) &= \frac{(\mu t)^{N_0-n}}{(N_0-n)!}e^{-\mu t} = \frac{(t/\tau)^{N_0-n}}{(N_0-n)!}e^{-t/\tau} \quad N_0 \geq n > 1, \\ P_0(t) &= 1 - \sum_{n=0}^{N_0-1}\frac{(\mu t)^n}{n!}e^{-\mu t} = 1 - \sum_{n=0}^{N_0-1}\frac{(t/\tau)^n}{n!}e^{-t/\tau} \end{aligned} \tag{22}$$

In Appendix B, we show a derivation of the solutions (22), (same as (24)). We now briefly present a description of the *sojourn time*. Which we use to theoretically *measure* the residence time in the radioactive decay phenomenon of our Poisson stochastic model.

4.3.1. Sojourn time

The *sojourn time (waiting time in any state)* is an intrinsic characteristic of any Markov process, where the change of state of the random variable is exponentially distributed (see [18] chap.16, and [33] chap.VI). In figure 1, we show a sketch of a Linear-**Pure Death Process** (**L-PDP**), (i.e., for $\mu_k = k\mu$). Likewise, we show the *sojourn times* in each stage and how the transitions follow an exponential distribution (decaying), which also has the *memoryless property*. Together we can apply to obtain the Poisson process as was done with the help of Eqs. (21), whose solution is (22).

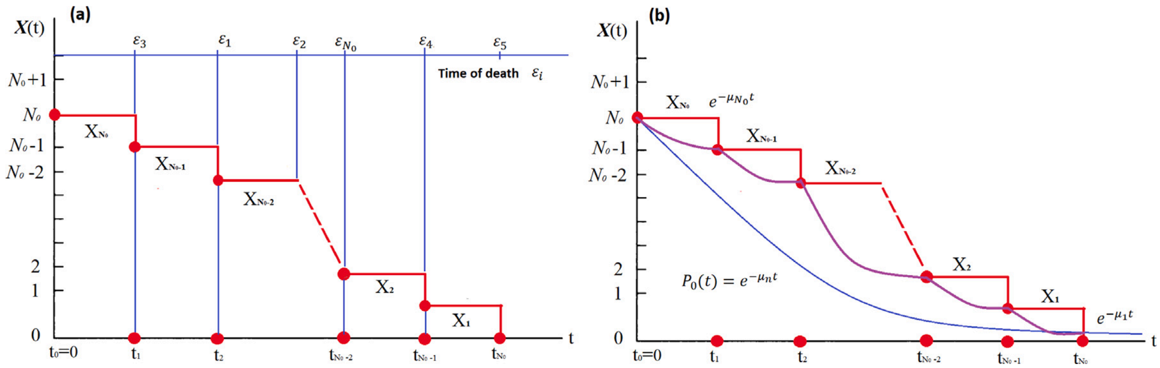

Fig. 1. Here we show the **Linear-Pure Death Process** (**L-PDP**): (a) Sketch of typical trajectories for a continuous-time Markov process (**CTMP**), in particular the pure death process **PDP**, which very accurately emulates a *Poisson-type decay process*, where the number of particles n decreases from $n = N_0$ at time $t = 0$ to $n = 0$ at $t \neq 0$. In addition, the **sojourns times** $X_n = \epsilon_n$ can be displayed, where $n = 0, 1, \dots, N_0 - 1, N_0$. (b) The curves represent that the transition probabilities are exponentially distributed with decay parameters μ_n the *death rate* ($\mu_n = n\mu$). For our system $\mu_1 = \mu_2 = \cdots = \mu_{N_0} = \mu = 1/\tau$, (without including a specific value of the physical decay parameter).

The sojourn time in state N_0 is denoted by X_{N_0}. Then, let $X(t)$ be a linear death process, denoting the number of survivors in the population at time t. With parameter $\mu_n = n\mu$ for $n = 0, 1, 2, \dots, N_0$. And let $\epsilon_1, \epsilon_2, \dots, \epsilon_{N_0}$, denote the death times of the members, individuals, or components of the cluster —the succession of death times should not necessarily have a chronological order of death, see Fig. 1-(a)—. X_{N_0} is the time of the earliest death or $X_n = min\{\epsilon_1, \epsilon_2, \dots, \epsilon_{N_0}\}$. Since the memoryless property of the exponential distribution holds and the lifetimes are independent, we have

$$\begin{aligned} P_r\{X_n > t\} &= P(min\{\epsilon_1, \epsilon_2, \dots, \epsilon_{N_0}\} > t) \\ &= P(\epsilon_1 > t, \epsilon_2 > t, \dots, \epsilon_{N_0} > t) \\ &= P(\epsilon_1 > t)P(\epsilon_2 > t)\cdots P(\epsilon_{N_0} > t) \\ P_r\{X_n > t\} &= \left[P(\epsilon_1 > t)\right]^{N_0} = e^{-N_0\mu t} \end{aligned} \tag{23}$$

That is, has an exponential distribution with parameter $N_0\mu$. Taking into account the sojourn time in the transition probability, the solution Eqs. (22) are transformed as

$$\begin{aligned} P_r\{X(t) = n\} = P_n(t) &= \frac{(N_0\mu t)^{N_0-n}}{(N_0-n)!}e^{-N_0\mu t} = \frac{(N_0 t/\tau)^{N_0-n}}{(N_0-n)!}e^{-N_0 t/\tau} \quad N_0 \geq n > 1, \\ P_r\{X(t) = 0\} = P_0(t) &= 1 - \sum_{n=0}^{N_0-1}\frac{(N_0\mu t)^n}{n!}e^{-N_0\mu t} = 1 - \sum_{n=0}^{N_0-1}\frac{(N_0 t/\tau)^n}{n!}e^{-N_0 t/\tau} \end{aligned} \tag{24}$$

This is the stochastic Poisson model based on a linear pure death process **(L-PDP)**, considering the sojourn time in each state of the system.

4.4. *Poisson's process as a pure birth process* ***(PBP)***

In Section 4.2 we show the Birth-Death process **(BDP)** modeled with Eqs. (8). Then, a **(BDP)** process is called a pure birth process **(PBP)** if $\lambda_i = \lambda$ in (8) for all $i \geq 0$ and $\mu_i = 0$ for all $i \geq 0$. It can be shown relatively easily that by solving Eqs. (8) with $N_0(0) = 0$, the *typical or standard Poisson distribution* (29) —coming from a **PBP**— is obtained as a solution (see [18,30,32,33]). We now apply the distribution to the radioactive decay process of nuclei —which can be visualized as in the toy model with multi-sided dice 2.1—. This Poisson model is to be universally used from the didactic teaching of the theory of probability and statistics. We can focus solely on its application to the phenomenon of radioactive decay, beginning with the perspective of radioactive dice.

In order to there to be an equivalence between the nuclei and the dice, it is necessary to increase, as already mentioned, the number of faces of each die involved, which represents each radioactive isotope. As a first case, it is shown that all the dice (nuclei) in decay as a whole are distributed according to a Poisson distribution (remembering that it must be fulfilled that the number of events is very large and the associated probability very small), because this distribution often provides a good model for the probability distribution of the X number of weird events that occur in space, time, or any other dimension, where the decay constant λ is a very sensitive parameter —together with the decay time period— to the values of each radioisotope and to the average value of r. v. X. Thus, we have for each dice toss, the result is subject to the random variable $X \sim f(x) = p(x) = \lambda^x e^{(-\lambda)}/x!$ (Poisson distribution). It can be easily demonstrated that its moment-generating function (MGF) is given by $M_X(t^*) = E[e^{(Xt^*)}] = e^{\lambda(e^{t^*}-1)}$, ($t^*$ does not represent time, it is a dummy parameter, as we defined it in paragraph 2.1). Now we have that for a set of n discrete random variables, i.e. $X_1, X_2, \dots, X_n$ with Poisson distribution each, the sum of all of them must result in another Poisson distribution

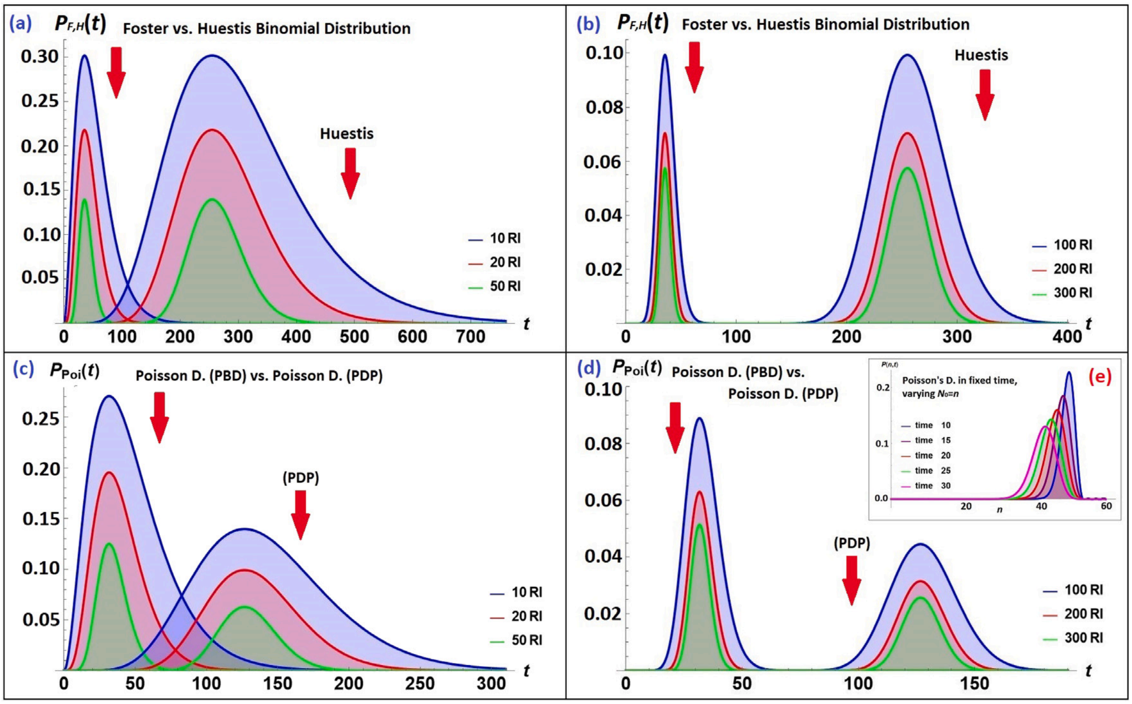

Fig. 2. In this figure, we show the plots corresponding to the four (stochastic) probability distribution models studied in this paper. In (a), we compare the model of Foster et al. with the Huestis model. The former utilizes the binomial PDF (4), exhibiting strong characteristics of a Pure Birth Process **(PBP)** (Yule Process). In contrast, Huestis's model employs the binomial PDF (2), showing strong features of the *Pure Death Process* (**PDP**). Both models operate on a relatively long timescale due to the small number of initial radioisotopes (**RIs**). In (b), we again compare the same models (4) vs. (2), but this time for a greater number of initial radioactive isotopes (**RIs**), resulting in a relatively short timescale due to the increased quantity of (**RIs**). In (c) and (d), we also compare our proposals based on the Poisson distribution versus the standard Poisson distribution. We juxtapose the *Pure Birth Process* **(PBP)** used as the standard in most statistical and probabilistic application studies with our proposed model, i.e., the *Pure Death Process* (**PDP**), utilizing the *sojourn time* which is fundamental in this process (see Sections Section 4.3, 4.3.1, Fig. 1, (22) and (24)). The effects on the time scales are similar to those observed for the binomial distribution process. Also, the figure (e) inset in (d), represents the Poisson distribution as a function of the number $N_0 = n$, for fixed times (dimensionless time t/τ) shown in the same figure. It should be noted that for large n the PDF fluctuates too much and suffers from overflow. For all plots we use *radioisotope*: ^{18}F with parameter (decay constant): $\mu = \lambda = 6.313 \times 10^{-3}$ min^{-1}. Initially 20% of isotopes decaying in all four combinations.
Source: Source parameter: [12,13].

—similar to the radioactive dice toy model in [4,9] shown in paragraph 2.1—, as we show below. Using the moment generating function, (see for example: [18,27,28,32,35–37]). We have $X_j \sim Poisson(X_j) = f(x_j)$, with $j = 1, 2, \ldots, n$; so we get

$$M_X(t^*) = E[e^{(t^* X_j)}] = e^{\lambda_j (e^{t^*} - 1)} \tag{25}$$

Adding all the random variables (with $\Sigma_n = \sum_{j=1}^n X_j$), we have

$$\begin{aligned} M_\Sigma(t^*) &= E[e^{(t^* \Sigma_n)}] = E[e^{(X_1+X_2+\cdots+X_n)t^*}] \\ &= \prod_{j=1}^n E[e^{(x_j t^*)}] = E[e^{(x_1 t^*)}]E[e^{(x_2 t^*)}] \cdots E[e^{(x_n t^*)}] \\ &= e^{\lambda_1 (e^{t^*}-1)} e^{\lambda_2 (e^{t^*}-1)} \cdots e^{\lambda_n (e^{t^*}-1)} \\ &= \prod_{j=1}^n \exp\left[\lambda_j (e^{t^*} - 1)\right] = \exp\left[\sum_{j=1}^n \lambda_j (e^{t^*} - 1)\right] \end{aligned} \tag{26}$$

This expression is once again a *Poisson distribution* for the n **r.v.**; with $\lambda_j = \langle x_j \rangle$ the *mean* of each **r.v.** (in this case, they are dimensionless parameters that represent the expected value). We can also add directly (without using the MGF method) two or more independent random variables (**i.r.v.**) as a *Poisson Process*, which should result in another *Poisson process*. In such a way that we have initially two r. v. X and Y. We are interested in the sum, i.e. $Z = X + Y$, where $Y = Z - X$, so we proceed to add both variables. Following this trend, we can generalize to the sum of n **i.r.v.** (that is, X_j, with $j = 1, 2, \ldots, n$) as a Poisson process. In both cases we obtain the following (in the Appendix A we make the complete demonstration of both formulations). For two independent and identically distributed random variables (**i.i.d.-r.v.**) $Z = X + Y$, where $X \sim Poisson(X, \lambda) = P_X(x)$, and $Y \sim Poisson(Y, \mu) = P_Y(y)$, and with help of *Discrete Convolution Formula*, we get $P_Z(z) = P_{X+Y}(z) = P_X(x) * P_Y(y) = \sum_x P_X(x) P_Y(y) = \sum_x P_X(x) P_Y(z - x)$ $\Rightarrow P_Z(z) = \frac{e^{-(\lambda+\mu)}}{z!} (\lambda + \mu)^z$.

So $Z = X + Y$ is Poisson distributed once again, and we just sum the parameters, which are the sum of means. If we make the change $\lambda \to \lambda_1$ and $\mu \to \lambda_2$, so that $Z \to S_2 = X_1 + X_2$, then the above equation takes the form

$$P_{S_2}(s_2) = P_{X_1+X_2}(s_2) = P_{X_1}(x_1) * P_{X_2}(x_2) = \frac{e^{-(\lambda_1+\lambda_2)}}{(s_2)!}\left(\lambda_1 + \lambda_2\right)^{s_2} \tag{27}$$

where $s_2 = x_1 + x_2$. But, what happens about a sum of more than two independent Poisson random variables? Suppose X_1, X_2, X_3 are random variables independent Poisson, and then $(X_1 + X_2)$ is then Poisson r.v., and then we can add on X_3 and still have a Poisson r.v. So $X_1 + X_2 + X_3$ must be a Poisson random variable. And so on, increasing to n i.r.v. of Poisson. In order to generalize, we have the n i.r.v., with $X_1, X_2, \dots, X_n$. Now, let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n i.r.v. of an independent trial process with common PDF (Poisson), defined on the integers. So, $S_n = (X_1 + X_2 + X_3 + X_4 \cdots) + X_n = S_{n-1} + X_n$. We have that i.r.v. are distributed by $X_1 \sim Poisson(X_1, \lambda_1) = P_{X_1}(x_1) \Rightarrow P(x_1) = \lambda_1^{x_1} e^{(-\lambda_1)}/(x_1!)$; $X_2 \sim Poisson(X_2, \lambda_2) = P_{X_2}(x_2) \Rightarrow P(x_2) = \lambda_2^{x_2} e^{(-\lambda_2)}/(x_2!); \dots \dots ; X_n \sim Poisson(X_n, \lambda_n) = P_{X_n}(x_n) \Rightarrow P(x_n) = \lambda_n^{x_n} e^{(-\lambda_n)}/(x_n!)$. Again, we use the *Discrete Convolution Formula* for several independent random variables,

$$\begin{aligned}
P_{S_n}(s_n) &= P_{X_1+X_2+\cdots+X_n}(s_n) = P_{X_1}(x_1) * P_{X_2}(x_2) * \cdots * P_{X_n}(x_n) \\
&= \sum_{\substack{x_1,\dots,x_n \\ s_n = x_1 + x_2 \cdots + x_n}}^{s_n} P_{X_1}(x_1) P_{X_2}(x_2) \cdots P_{X_n}(x_n) \\
&= \sum_{\substack{x_1,\dots,x_n \\ s_n = x_1 + x_2 \cdots + x_n}}^{s_n} \frac{\lambda_1^{x_1} e^{-\lambda_1}}{(x_1!)} \frac{\lambda_2^{x_2} e^{-\lambda_2}}{(x_2!)} \cdots \frac{\lambda_n^{x_n} e^{-\lambda_n}}{(x_n!)} \\
&= \frac{e^{-\sum_{j=1}^{n} \lambda_j}}{(s_n!)} \left(\lambda_1 + \lambda_2 + \lambda_3 + \cdots \lambda_n\right)^{s_n} \\
&= \frac{e^{-\sum_{j=1}^{n} \lambda_j}}{(s_n!)} \left(\sum_{j=1}^{n} \lambda_j\right)^{s_n}
\end{aligned} \tag{28}$$

This is the *probability mass function (p.m.f.)* for n random variables distributed as Poisson process model. Where $S_n = X_1 + X_2 + X_3 \cdots + X_{n-1} + X_n = S_{n-1} + X_n$, and the parameters $\lambda_j = \langle x_j \rangle$ are the mean (expected value). Therefore $\sum_{j=1}^{n} \lambda_j = \sum_{j=1}^{n} \langle x_j \rangle = \langle s_n \rangle$ are simply the sum of expected values. Thus, we can write in a more compact form Eq. (28), which coincides with the known most basic Poisson distribution.

$$P_{S_n}(s_n) = \frac{e^{-\langle s_n \rangle}}{s_n!} \langle s_n \rangle^{s_n} = \left(\frac{\langle s_n \rangle^{s_n}}{s_n!}\right) e^{-\langle s_n \rangle} \tag{29}$$

Thus, based on a standard time-dependent Poisson process model (see [18,29,30,32,33]) this is expressed as

$$P_{X=n}(t; \lambda) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Valid for a single event (decay). However for the case of the sum of n events, where we have the total value of nuclei are $n = N_0$ and $S_n = n_1 + X_2 + X_3 \cdots X_n \to X$, for simplicity and to get rid of sub-indexes. The value of the parameter or decay constant depends on each substance (radioisotopes) in each event, which would have the same value for all the nuclei involved (unless otherwise stated), and it would be $\lambda_1 = \lambda_2 = \cdots \lambda_n = \lambda$. In addition, the process it depends on time, i.e. $\lambda \to \lambda t$.[5] So, $\sum_{j=1}^{n=N_0} (\lambda_j t) = N_0 \lambda t$, so we obtain

$$P_n(t) = \left[\frac{(N_0 \lambda t)^n}{n!}\right] e^{-N_0 \lambda t} \tag{30}$$

Which similarly includes sojourn time as in (24), (see Section 4.3.1).

4.4.1. Partial nuclear decay : Prior and conditional probability

To get a clearer insight of the radioactive decay process, let's make an arbitrary partition (for example: assuming that $X_m < X_n$, from the initial population, divide the entire population: N_0 into strata or subpopulations $N_0 \gg x$ and so on) in the total amount of radioactive isotopes (nuclei or atoms), considering that at the beginning of the reaction they are divided into two groups (those that undecayed immediately in the first period of time, and those that do decayed in the next period of time, i.e. the elapsed time between the $(n-1)$st and the nth event.) That is, one of the groups decays faster than the other (the smaller group to be precise), as usually happens in reality, since not all the nuclei (isotopes) involved decay at the same time. For this we consider two sums (one for each group) of the total isotopes (N_0), where one can be much larger than the other, initially being also independent Poisson random variables. Then, we first calculate the prior probability using the result of Eq. (28) as follows: the sum is rewritten in two terms for both groups, then the probability distribution for the total of the two groups is obtained, that is

$$P_{S_n}(s_n) = \frac{e^{-\sum_{j=1}^{n} \lambda_j}}{(s_n!)} \left(\sum_{j=1}^{n} \lambda_j\right)^{s_n}$$

[5] Actually, this is justified by introducing the **sojourn time** seen in Section 4.3.1.

$$
\begin{aligned}
&= \frac{e^{-\sum_{j=1}^{m}\lambda_j}}{(s_m!)}\frac{e^{-\sum_{j=m+1}^{n}\lambda_j}}{(s_{m+1}!)}\left[\sum_{j=1}^{m}\lambda_j+\sum_{j=m+1}^{n}\lambda_j\right]^{s_n}\\
&= \frac{e^{-\Lambda_m}}{(s_m!)}\frac{e^{-\Lambda_n}}{(s_{m+1}!)}\left[\Lambda_m+\Lambda_n\right]^{s_n} = \frac{e^{-\Lambda_m}}{(s_m!)}\frac{e^{-\Lambda_n}}{(s_{m+1}!)}\sum_{k=0}^{s_n}\binom{s_n}{k}\Lambda_m^{s_n-k}\Lambda_m^k\\
&= \frac{e^{-(\Lambda_m+\Lambda_n)}}{(s_m!)(s_{m+1}!)}\sum_{k=0}^{s_n}\binom{s_n}{k}\Lambda_m^{s_n-k}\Lambda_m^k
\end{aligned}
\tag{31}
$$

Which is equivalent to Eq. (28), where we simply split the total amount of isotopes into two groups (arbitrarily). Where, $\Lambda_m = \sum_{j=1}^{m}\lambda_j$ and $\Lambda_n = \sum_{j=m+1}^{n}\lambda_j$; $s_m! = x_1!x_2!\cdots x_m!$ and $s_{m+1}! = x_{m+1}!x_{m+2}!\cdots x_n!$. If considered as a stochastic Poisson model, of the type of equations Eqs. (28) and (30), we obtain a split version of the prior probability factored, but not completely separable due to the factorials in the denominator, that is

$$
P_X(t;\lambda;N_0) = \frac{(\Lambda_m\lambda t+\Lambda_n\lambda t)^x}{(s_m!)(s_{m+1}!)}e^{-\Lambda_m\lambda t}e^{-\Lambda_n\lambda t}
$$

where $N_0 = N_{01}+N_{02} = \Lambda_m+\Lambda_n$, $(\Lambda_m<\Lambda_n)$ the partition in the two groups above. This equation is also equivalent to (30).

4.4.2. Conditional probability, and conditional expected value

Now, we calculate the conditional probability and the conditional expected value, similarly to the one obtained in Eq. (31). That is, as mentioned above, we split (arbitrarily) our sample space into two partitions: Let $S_n = X_1+X_2+\cdots+X_m+X_{m+1}+\cdots+X_n = \Lambda_m+\Lambda_n$ be the same sum of n independent Poisson random variables. Thereby, the conditional probability is

$$
\begin{aligned}
P_{S_n}[X_m = s_m|X_m+X_n = s_n] &= \frac{P[X_m = s_m, X_m+X_n = s_n]}{P[X_m+X_n = s_n]}\\
&= \frac{P[X_m = s_m, X_n = s_n-s_m]}{P[X_m+X_n = s_n]}\\
&= \frac{P[X_m = s_m]P[X_n = s_n-s_m]}{P[X_m+X_n = s_n]}\\
&= \frac{e^{-\Lambda_m}\Lambda_m^{s_m}}{(s_m!)}\frac{e^{-\Lambda_n}\Lambda_n^{(s_n-s_m)}}{(s_n-s_m)!}\left[\frac{e^{-(\Lambda_m+\Lambda_n)}}{(s_n!)}\right]^{-1}\\
&= \frac{(s_n!)}{(s_n-s_m)!(s_m!)}\Lambda_m^{s_m}\Lambda_n^{(s_n-s_m)}\left(\frac{\Lambda_m+\Lambda_n}{\Lambda_m+\Lambda_n}\right)^{s_m}\\
&= \binom{s_n}{s_m}\left(\frac{\Lambda_m}{\Lambda_m+\Lambda_n}\right)^{s_m}\left(\frac{\Lambda_n}{\Lambda_m+\Lambda_n}\right)^{s_n-s_m}
\end{aligned}
\tag{32}
$$

That is, the conditional distribution of $X_m = s_m$ given that $X_m+X_n = s_n$ (Assuming that $X_m < X_n$) is the *Binomial Distribution* with parameters s_n and $\Lambda_m/(\Lambda_m+\Lambda_n)$. From this, it is easy to notice that the expected value for this *new binomial distribution* is given by:

$$
\mathbf{E}[X_m|X_m+X_n = s_n] = s_n\left(\frac{\Lambda_m}{\Lambda_m+\Lambda_n}\right)
\tag{33}
$$

We can notice that when comparing the Eqs. (32) and (33) with Eq. (4) and its mean, and making the following changes $\Lambda_n \to N_{01}\lambda t$ and $\Lambda_m \to N_{02}\lambda t$, (with $N_0 = N_{01}+N_{02}$, i.e. the total number of nuclei remains constant), we get that: $\Lambda_m/(\Lambda_m+\Lambda_n) = \frac{N_{02}}{N_0} = (1-e^{-\lambda t}) = p(t)$, the decay probability, and $\Lambda_n/(\Lambda_m+\Lambda_n) = \frac{N_{01}}{N_0} = e^{-\lambda t} = [1-p(t)] = q(t)$, the undecayed probability. We want to argue that

$$
\begin{aligned}
N_{01} &= N_0e^{-\lambda t}\\
N_{02} &= N_0(1-e^{-\lambda t})
\end{aligned}
\tag{34}
$$

Also with mean $\mathbf{E}(X) = N_0(1-e^{-\lambda t})$, and $\mathbf{Var}(X) = N_0e^{-\lambda t}(1-e^{-\lambda t})$. Eqs. (34), prove their validity because they are similar to those obtained in the models with the Eqs. (1), (2), and (4). It is worth noting that these equations also allow us to see that the probability of decay of radioisotopes is updated at each time interval; because it is an exponential stochastic process (i.e. a Markov chain as a process memoryless).

In summary, the previous results in Eqs. (32), (33), and (34), based on the so-called *pure birth process* **(PBP)** (of Poisson and Binomial), allow us to observe that in reality the behavior and evolution in the radioactive decay of a isotopes system —being conditioned to partial decay for different periods of time— this will follow a binomial distribution, such a distribution comes from assuming that initially, it is an independent Poisson random variable and that not all isotopes decay simultaneously. This shows the trend of isotopes radioactive decay to follow a binomial distribution when the decay is initially conditioned, splitting the total amount of radioisotopes (as mentioned above, those not yet decaying and those beginning to decay). We should also keep in mind that this is a random physical phenomenon in which undecayed isotopes of a first type (say A with N_{01} in Eq. (34)), decay to a second type of radioisotopes (say B with N_{02} in Eq. (34)); see a more physical-deterministic treatment in the classic treatise by Kenneth S. Krane: *Introductory Nuclear Physics*, ([11] chapter 6), where we obtains equations similar to (34).

5. Fluctuations and entropy

In this section, we show the latest original theoretical results of our study, namely, the intrinsic fluctuations of the decay process and the entropy associated with the same physical phenomenon, which as far as we know there are no previous studies. The fluctuations are calculated through the *Fano factor*, and we derive the *Shannon entropy metric* as a stochastic process for the decay phenomenon.

5.1. Relative fluctuations (*RF*): Fano factor

Every system with a random nature inevitably has intrinsic fluctuations that can be accounted for using different statistical variability metrics. Fluctuations refer to variations or changes in a system or process over time. In this section, we analyze the fluctuations through the Fano Factor, namely, for the binomial and Poisson distribution in the context of the birth–death processes, which have significant differences as we have seen in the previous sections 3.2, 4.3 and 4.4.

In section[6] 4.4, we have shown the prior and conditional probabilities, respectively. So, we can deduce and conclude that radioactive decay behaves and evolves in two stages. (both distributions) according to different features such as: ***(i)*** First, since the binomial distribution can approximate the Poisson distribution, then, they intrinsically have a connection —at least mathematically—, where the main requirement for the binomial is that the number of trials or items —in this case the radioisotopes— are large enough ($S_n = N_0 \to \infty$), and the assigned probabilities are very small ($p(t) \to 0$), because $N_0 p(t) < \infty$. This last requirement is equivalent to having the number of nuclei that decay are much smaller than the number of radioactive nuclei that are *waiting — sojourn time* 4.3.1, which is more relevant than is believed — for the disintegration (i.e. $x \ll N_0$).

(ii) Second, the time intervals (time period), the parameters (decay constant) and the decay lifetime ratio (half-lives) for each specific radiotracer are very sensitive as they evolve. ***(iii)*** And third, if we consider the *relative fluctuations f_r or Factor Fano.*[7] This factor is widely used in particle counting statistics —photons, electro-photons, with high and low energy— in detection processes; also in studies on fluctuations in the noise of electro-optical signals, and neuroscience (see for example: [22,38,39]). In mathematical statistics it is known as the *Coefficient of Variation*[8]: $CV = \sigma/\mu$. In each stage of disintegration as we argue in (*i*) and (*ii*), when we have a large number of radio-isotopes undecayed, these on average behave deterministically. However, those that decay (at a constant rate), which are much less, have a more random behavior.

5.1.1. RF for binomial and Poisson distribution as a pure birth process **(PBP)**

Relative fluctuations can be calculated through the ratio between the standard deviation and the mean of the distribution. That is, according to the binomial distribution, as the number of elements in a system increases, the mean value $\bar{x} = E(X) = \mu_X$ increases linearly with $n = N_0$, whereas the standard deviation $\sigma = (Var(X))^{1/2}$ increases only as the square root of N_0, i.e.

$$f_r = \frac{\sqrt{Var(X)}}{\mathbf{E}(X)} = \sqrt{\frac{N_0 e^{-\lambda t}(1-e^{-\lambda t})}{(N_0(1-e^{-\lambda t}))^2}} = \frac{1}{\sqrt{N_0}}\sqrt{\frac{e^{-\lambda t}}{(1-e^{-\lambda t})}} \tag{35}$$

Making a Taylor series expansion, we obtain

$$\begin{aligned}\frac{\sqrt{Var(X)}}{\mathbf{E}(X)} &= \frac{1}{\sqrt{N_0}}\left\{\frac{1}{\sqrt{\lambda t}} - \frac{\sqrt{\lambda t}}{4} + \frac{1}{196}\lambda^{3/2}t^{3/2} + \frac{1}{384}\lambda^{5/2}t^{5/2}\right. \\ &\quad \left. - \frac{1}{10240}\lambda^{7/2}t^{7/2} + \frac{19}{368640}\lambda^{9/2}t^{9/2} + \cdots + O(t)^{n/2}\right\}\end{aligned}$$

Finally, by discarding higher order terms in the expansion, we can approximate the relative fluctuations, consequently

$$f_r = \frac{\sqrt{Var(X)}}{\mathbf{E}(X)} \approx \frac{1}{\sqrt{N_0 \lambda t}} \tag{36}$$

we show that the relative fluctuations are very small if N_0 is very large (deterministic tendency). However, if N_0 is small, the time and the parameter λ influence much more the random behavior in the decay; as usually happens in these cases. From Eq. (36)

---

[6] In this section we use the λ parameter interchangeably with the stochastic process, however, they can be used with other parameters such as μ and γ. Mathematically, it does not affect the result.

[7] Fano factor [38] is a measure of variability of a counting process defined as

$$F(s) = \frac{Var[N(s)]}{E[N(s)]}, \quad s > 0$$

where $N(s)$ is a equilibrium counting (stochastic) process that describes the number of items —radioisotopes in this case— in an interval $(0, s]$, $s > 0$, where time zero is randomly placed with respect to the sequence of items (radioisotopes)

[8] The *coefficient of variation (CV) or Fractional standard deviation* is a statistical measure of the spread of data points in a data series around the mean. The coefficient of variation represents the relationship (ratio) between the standard deviation and the mean, and is a useful statistic for comparing the degree of variation in one data series with another, even if the means are drastically different from each other.

In the context of statistical physics, fluctuations are defined as $f_r = \sigma_X/\mu_X = \frac{\sqrt{\langle X^2\rangle - \langle X\rangle^2}}{\langle X\rangle} = \frac{1}{\sqrt{\langle X\rangle}}\sqrt{1+\langle X\rangle}$, known as the *relative root-mean-square (rms)* (without the approximation) we can see that $Var(X) = \mu_X e^{-\lambda t}$, i.e., is proportional to μ_X, the expected value or mean. And the relative fluctuation: $f_r = \left(1/\sqrt{\mu_X}\right) e^{-\lambda t/2}$, i.e., is proportional to the inverse of root mean square (see footnote 7 and 8).

For the *Poisson (process) distribution*, (we do not need an approximation) we directly obtain

$$f_r = \frac{\sqrt{Var(X)}}{\mathbf{E}(X)} = \sqrt{\frac{N_0 \lambda t}{(N_0 \lambda t)^2}} = \frac{1}{\sqrt{N_0 \lambda t}} \tag{37}$$

Both approaches with the Poisson and Binomial distributions are mathematically complementary, even in a stochastic pure birth process context 4.4. Nevertheless, as we will see below 5.1.2 in the context of a stochastic pure death process, there are marked differences. From the Eqs. (36) and (37) — see Fig. 3 with the plots of Eqs. (35) and (37) without the approximation (36), which is practically equal to (37)— we show that the relative fluctuation between both distributions (stochastic processes) are practically the same. Which means that on average the variability of radioactive decay between one distribution and another evolve statistically in a similar way. However, these distributions have particular mathematical characteristics. As they are continuous functions over time (stochastic pure birth process), their graphic performance shows significant differences in their evolution (maximum values, for example) relative to the calculated probabilities. Remember that the binomial distribution is proportional to a factorial coefficient —total number of samples and trials— and the Poisson distribution is proportional to the ratio between the mean and factorial of the number of trials. For this reason it is lower than the binomial, with the same values. All this depends on the conditions mentioned above —***(i)***, ***(ii)*** and ***(i)***— where the stochastic process, i.e., the random phenomenon is updated in each time interval as we show in Section 4.4.1, and where the *sojourn time* plays a very important role 4.3.1. That is, if a radioisotope randomly decays, its decay rate converges with the rate of the exponential type that has the lowest decay rate. Because the longer an isotope takes to decay, the more likely it is to be a type of radioactive element with the slowest decay rate.

Being a theoretical study, we assume ideal conditions, but we are aided by at least one real physical parameter. On the other hand, the Fano factor that measures the fluctuations is an ideal case, since if we consider a more realistic situation we should include (adjust) noise variations due to excitation events, which alter the coherence broadening in the signal. (see for example [40]). Whether the experimental detection process is considered will depend on the detector used and its efficiency: Such as the physical and technological features —parameters, materials, radioactive substances—, time periods, photo-detection techniques and/or detection of alpha, beta, gamma particles (or radiation), so on. Which is outside the scope of this work —for a more detailed study, those interested can consult the references: [10,13,19,20,22,23]—.

5.1.2. RF for binomial (and Poisson) distribution as a pure death process ***(PDP)***

Here show the fluctuations for **(PDP)**

$$f_r = \frac{\sqrt{Var(X)}}{\mathbf{E}(X)} = \sqrt{\frac{N_0 e^{-\lambda t}(1 - e^{-\lambda t})}{(N_0 e^{-\lambda t})^2}} = \frac{1}{\sqrt{N_0}} \sqrt{e^{\lambda t} - 1} = \frac{e^{\lambda t/2}}{\sqrt{N_0}} \sqrt{1 - e^{-\lambda t}} \tag{38}$$

Similarly to Eqs. (35) and (36) we making a Taylor series expansion:

$$\begin{aligned} \frac{\sqrt{Var(X)}}{\mathbf{E}(X)} = \frac{1}{\sqrt{N_0}} \Bigg\{ & \sqrt{\lambda t} + \frac{1}{4}\lambda^{3/2} t^{3/2} + \frac{5}{96}\lambda^{5/2} t^{5/2} + \frac{1}{128}\lambda^{7/2} t^{7/2} \\ & - \frac{79}{92160}\lambda^{9/2} t^{9/2} + \frac{3}{40960}\lambda^{11/2} t^{11/2} + \frac{71}{12386304}\lambda^{13/2} t^{13/2} + \cdots + O(t)^{n/2} \Bigg\} \end{aligned}$$

We can discard higher-order terms, considering that the parameters and the quotients tend quickly to zero.

$$f_r = \frac{\sqrt{Var(X)}}{\mathbf{E}(X)} \approx \frac{\sqrt{\lambda t}}{\sqrt{N_0}} \tag{39}$$

We can observe that now the fluctuations (Fano factor) for the binomial process —as a pure death process— are proportional to the direct square root of the product of the time by the parameter, unlike the fluctuations for the binomial process —a pure birth process— in Eq. (36), where this process is proportional to the inverse of the square root. Regarding the Poisson process, in both cases, the fluctuations vary in the same way (see Eq. (37), and figure 3).

5.2. Stochastic entropy

Here we show entropy metrics in the mathematical context of *Shannon's information theory*. We compute the entropy function for the standard Poisson and binomial stochastic processes. Seldom used in a physical context. And as far as we know, it has never been used for these distributions —binomial and Poisson-like— in radioactive decay applications, as in the case of our study. We perform the analysis purely theoretically, leaving for future studies a deeper research into applications in other physical contexts. *Shannon's entropy* is a mathematical measure of the amount of uncertainty or randomness in a given set of data or information. It was introduced by Claude Shannon in his seminal paper *A Mathematical Theory of Communication* in 1948 [21]. Shannon's entropy is often used in the context of information theory to quantify the amount of information contained in a message, signal or data

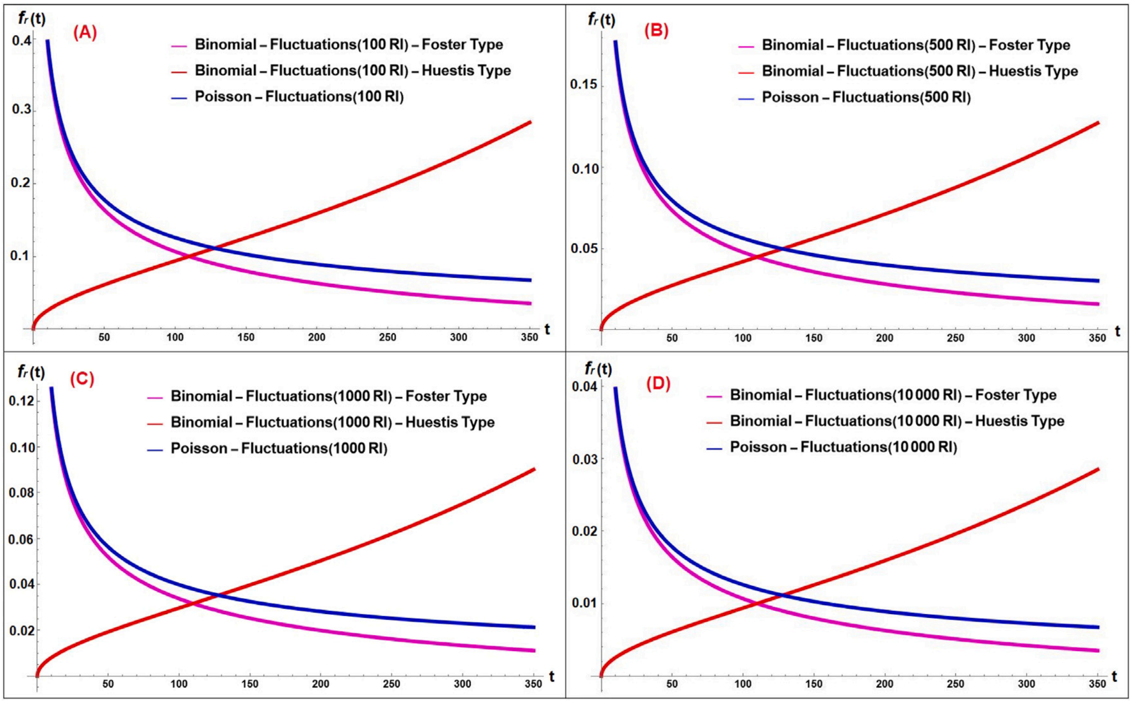

Fig. 3. *Binomial (Huestis vs. Foster et al.) versus Poisson Fluctuations* (*Fano factor*): In this figure we can visualize the plots (**A** to **D**) of the Eqs. (35), (37), and (38) of the Fano factor corresponding to the fluctuations of the Binomial (Huestis vs. Foster et al.) and Poisson process, for the parameter: $\mu = \lambda = 6.313 \times 10^{-3}$ min^{-1}. However, these processes are matched by approximating through of parameters —i.e. the decay constants, and analytically doing a Taylor series expansion — in a limiting case obtained in Eq. (36). In all cases (**A** to **D**), the binomial process (Foster et al. case (**PBP**)) *decays* or decreases its fluctuations faster than the Poisson process, nevertheless, both processes fluctuate similarly in the equation approximation (36). But the binomial process decreases faster. Regarding the case of the Huestis-type binomial process (**PDP**), the Fano factor indicates that the fluctuations increase with time.

stream —for more information and description of this theory, see the excellent treatise by Cover and Thomas: [41]—. Shannon's entropy is defined as:

$$\mathbf{H}(X) = -\sum_{k=1}^{\infty} P_X(k) \log P_X(k) = -\mathbf{E}\Big[\log P_X(k)\Big] \tag{40}$$

where $\mathbf{H}$[9] is the entropy of the system, $P_X(x)$ is the probability of a particular outcome x, and $\log(\cdot)$ is the binary logarithm. The entropy is measured in **bits** if the base of the logarithm is 2. If the base of the logarithm is the number **e** (natural logarithm, inverse of the exponential function), entropy is measured in **nats**. In this case, the entropy value represents the average number of **nats** needed to represent or encode each event. The formula can be interpreted as the *expected value* of the information content of a message, weighted by its probability of occurrence. It measures the amount of uncertainty or randomness in the system: the higher the entropy, the more uncertain the system is. It should be noted that entropy is a relative measure and does not have an absolute scale. The magnitude of entropy depends on the probabilities (or probability density) of the events in the system or data set being analyzed. Higher entropy values indicate greater uncertainty or randomness, while lower entropy values indicate more order or predictability. Shannon's entropy has found many applications in various fields, including cryptography, data compression, signal processing, statistical mechanics, and now in radioactive decay. It provides a powerful tool for analyzing the properties of complex systems and understanding the fundamental limits of information processing.

5.2.1. *Poisson stochastic entropy*

This paper theoretically demonstrates the entropy function in binomial and Poisson stochastic processes context used for radioactive decay. It should be mentioned that the entropy for these distributions has already been calculated by other authors (among others) in purely mathematical contexts by M. Cheraghchi [43] with exhaustive generalizations and, on the other hand, with an approach directed to physical optics by A. Martínez [44]. We think they are among the most outstanding and we follow their trend. However, our focus is specifically on stochastic processes, where to our knowledge the entropy of the mentioned distributions

[9] In the classical statistical mechanics context [42], the Gibbs (and Boltzmann) entropy is traditionally represented by $S = -k \sum_{i=1} p_i \ln p_i$ ($k = cte. = 1$). But when the energy remains fixed, and all the states are equally accessible —that is, with probability $p_i = p = 1/w$ substituted in the Gibbs entropy—, the so-called Boltzmann entropy is obtained: $S = \ln w$.

has not been calculated. Now, using the definition of entropy (40) with the aforementioned stochastic distributions (see (4) for standard binomial process, and (29) to (30) for standard Poisson process), we proceed to show the respective entropies: (with $\langle s_n \rangle = s$)

$$
\begin{aligned}
\mathbf{H}(X_{Poisson}) &= \mathbf{H}_{Poiss} = -\sum_{k=0}^{\infty}\left(\frac{s^k}{k!}e^{-s}\right)\log\left(\frac{s^k}{k!}e^{-s}\right) \\
&= -\sum_{k=0}^{\infty}\frac{s^k}{k!}e^{-s}\left[k\log(s)-\log(k!)-s\right] \\
&= -e^{-s}\left\{\sum_{k=0}^{\infty}\frac{s^k}{k!}k\log(s)\right\} + se^{-s}\underbrace{\left[\sum_{k=0}^{\infty}\frac{s^k}{k!}\right]}_{e^s} + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\left[\log(k!)\right] \\
&= \left(s - s\log(s)\right) + \sum_{k=0}^{\infty}\frac{s^k}{k!}e^{-s}\left[\log\Gamma(k+1)\right] \\
&= \left(s - s\log(s)\right) + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\int_0^{\infty}\left[k-\frac{1-e^{-kx}}{1-e^{-x}}\right]\frac{e^{-x}}{x}dx \\
\mathbf{H}_{Poiss} &= \left(s - s\log(s)\right) + \int_0^1\left[\frac{1-e^{-s(1-u)}}{1-u}-s\right]\frac{du}{\log(u)}
\end{aligned}
\tag{41}
$$

where we use the gamma function ($\Gamma(z+1)=z!$) to express the factorial term. And the $\log(\cdot)$ function can be in binary base or the natural logarithm base, i.e. in base $\log_e(\cdot)=\ln(\cdot)$. The integral in formula (41) is proper and converges, nevertheless, note that in extreme limits (i.e., at $u=0$ and $u=1$), it may have convergence problems and behaves as follows [44]: At $u=0$ it tends to zero, and at $u=1$, we obtain the convergence by doing a Taylor series expansion on the exponential function

$$
\lim_{u\to 1}\left[\frac{1-e^{-s(1-u)}}{1-u}-s\right]\frac{1}{\log(u)} = \frac{s^s}{2}
$$

Expressions in terms of factorials and logarithmic functions are expressed with the help of the Gamma function in its different forms.

$$
\begin{aligned}
\log(k!) &= \log\Gamma(k+1) = \int_0^{\infty}\left[k-\frac{1-e^{-kx}}{1-e^{-x}}\right]\frac{e^{-x}}{x}dx \\
\log\Gamma(k+1) &= \int_0^1\left[\frac{1-u^k}{1-u}-k\right]\frac{du}{\log(u)}.
\end{aligned}
\tag{42}
$$

The first expression, in (42), is obtained by means of Ref. [45] (formula (9), page 360); where $u=e^{-x}$, —see Appendix C for more details—. We do not make any approximation for the factorial expression nor the logarithm function with the factorial, we use the Gamma function and exact integral representations. Because it has become customary in the physical literature to make use of Stirling's formula approximation. Here, we get, similarly to the cited Refs. [43,44], analytical, general, and exact expressions. Our approach, unlike the previous uses, is as we have already mentioned to obtain the Entropy of binomial and Poisson stochastic processes. Our analytical route is similar, but not the same as that of the cited authors, in Appendix C, we show these formulas using our own approach. The reason for obtaining an exact analytic formula is that its numerical evaluation —with the use of computers and numerical algorithms for the integral— is easier and more accurate than using the Stirling approximation.

Now we represent the formula as a function of time, considering that it is a stochastic process. From this, we finally obtain the entropy for the Poisson process. Making the following changes $s \to N_0\lambda t$ and $\log(u) \to \log_e(u) = \ln(u)$, where u should be considered an auxiliary variable.

$$
\mathbf{H}_{Poiss}(X(t)) = \left(N_0\lambda t - (N_0\lambda t)\ln(N_0\lambda t)\right) + \int_0^1\left[\frac{1-e^{-N_0\lambda t(1-u)}}{1-u}-N_0\lambda t\right]\frac{du}{\ln(u)}
\tag{43}
$$

Remark: We describe of formulas (41) and (43) to give us a better insight into their application. Formula (41) is similar to that obtained by the cited authors [43,44]. But, we developed different procedures to derive it; however, as far as we know, the formula (43) is unpublished. The application of these formulas depends on the specific stochastic context of interest. Formula (41) emphasizes using statistical and/or physical parameters, not considering how it evolves over time. In formula (43), time is explicitly included. Remember that we are dealing with a stochastic process of a discrete random variable (**d.r.v.**) in continuous time (*sojourn time*). Nevertheless, caution must be taken when using this formula, where we must take the time (individual, interval, or average) for each *stochastic trajectory*,[10] setting the total number $n = N_0$. In such a way that a family of trajectories is formed (*ensemble*). In order to calculate the integral, not only because it is numerical, but because each trajectory must be fixed in a specific time interval. One way to validate the entropy calculated for this type of stochastic process —Poisson and the binomial— is to look for a bound that

[10] A ***stochastic trajectory*** refers to the path or sequence of events followed by a system or process that is subject to stochastic (random or probabilistic) influences. In other words, it describes the evolution or behavior of a system over time when there is inherent uncertainty or randomness involved.

compares its plots, approximately. Likewise, one approach is to carry out a comparison using Gaussian Maximum Entropy **GME**[11] (see: [41], chapters 8, and 12. Also arXiv-paper [46]) —which is derived from the differential entropy of the Normal distribution, of a continuous random variable (**c.r.v.**) at continuous time— which works as an upper bound for the entropy (stochastic) of Poisson. We plot the numerical simulation of (41) in Fig. 4(**a**). And the binary type term $\left(s - s\log(s)\right)$ in Fig. 4(**b**). The **GME** comparison in Fig. 4(**c**).

5.2.2. Binomial stochastic entropy

Similarly, for binomial stochastic process, we can get the entropy function, i.e., following someone steps of Ref. [43], we calculate this entropy, the details we send to Appendix D.

$$\begin{aligned}
\mathbf{H}(X_{Binomial}) &= \mathbf{H}_{Bin} = -\sum_{k=0}^{\infty}\binom{n}{k}p^k\left(1-p\right)^{n-k}\log\left[\binom{n}{k}p^k\left(1-p\right)^{n-k}\right] \\
&= -\sum_{k=0}^{\infty}\frac{\Gamma(n+1)p^k\left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)}\log\left[\frac{\Gamma(n+1)p^k\left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)}\right] \\
&= -\sum_{k=0}^{\infty}\frac{\Gamma(n+1)p^k\left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)}\times \\
&\times\left[\log\Gamma(n+1)-\log\Gamma(k+1)-\log\Gamma(n-k+1)-k\log p+(n-k)\log\left[1-p\right]\right] \\
&= nh(p)-\log\Gamma(n+1)+\sum_{k=0}^{\infty}\log\Gamma(k+1)\binom{n}{k}p^k\left(1-p\right)^{n-k}+\sum_{k=0}^{\infty}\log\Gamma(n-k+1)\binom{n}{k}p^k\left(1-p\right)^{n-k} \\
\mathbf{H}_{Bin} &= nh(p)-\log\Gamma(n+1)+\mathbf{E}\left[\log\Gamma(k+1)\right]+\mathbf{E}\left[\log\Gamma(n-k+1)\right]
\end{aligned} \tag{44}$$

where

$$h(p) = -p\log p-(1-p)\log(1-p) \tag{45}$$

This is the binary entropy (we can see more details in Appendix D). And ($P_n(k) = f_{Bin}(k) = B(k;n,p) = \binom{n}{k}p^k(1-p)^{n-k}$ is the binomial distribution) the expected values are

$$\mathbf{E}\left[\log\Gamma(k+1)\right] = E_{Bin}(n,p) = \sum_{k=0}^{\infty}\log\Gamma(k+1)f_{Bin}(k) \tag{46}$$

$$\mathbf{E}\left[\log\Gamma(n-k+1)\right] = E_{Bin}(n,1-p) = \sum_{k=0}^{\infty}\log\Gamma(n-k+1)f_{Bin}(k) \tag{47}$$

After several transformations and using the previous results ((44) to (47)), we can obtain the entropy function for the binomial distribution, i.e.,

$$\begin{aligned}
\mathbf{H}_{Bin}(X) &= nh(p)-\log\Gamma(n+1)+E_{Bin}(n,p)+E_{Bin}(n,1-p) \\
&= nh(p)-\int_0^1\left[\frac{(1-up)^n}{\log(1-u)}\right]du-\int_0^1\left[\frac{\left(1-u(1-p)\right)^n}{\log(1-u)}\right]du+\int_0^1\frac{1+(1-u)^n}{u\log(1-u)}du \\
\mathbf{H}_{Bin}(X) &= nh(p)+\mathbf{H}_1+\mathbf{H}_2-\mathbf{H}_3
\end{aligned} \tag{48}$$

These last $\mathbf{H}_j$ functions are developed and well expressed in Appendix D. Which are more convenient to understand the scope and limitations of binomial entropy. Thus, we obtain the stochastic entropy for the binomial distribution, using the functions $p(t) = (1-e^{-\lambda t})$ and $1-p(t) = e^{-\lambda t}$ (Foster et al. type, similarly can be done for the Huestis type distribution; see Section 3.2), that is,

$$\mathbf{H}_{Bin}(X(t)) = nh(p(t))-\int_0^1\frac{\left[1-u(1-e^{-\lambda t})\right]^n}{\log(1-u)}du-\int_0^1\frac{\left[1-ue^{-\lambda t}\right]^n}{\log(1-u)}du+\int_0^1\frac{1+(1-u)^n}{u\log(1-u)}du \tag{49}$$

As in the case of the Poisson process, we can also make the change in the base of the logarithm, i.e., $\log(u)\to\log_e(u) = \ln(u)$ (and remember that $n = N_0$). Finally, we get the *binomial stochastic process entropy function*

$$\mathbf{H}_{Bin}(X(t)) = nh(p(t))-\int_0^1\frac{\left[1-u(1-e^{-\lambda t})\right]^n}{\ln(1-u)}du-\int_0^1\frac{\left[1-ue^{-\lambda t}\right]^n}{\ln(1-u)}du+\int_0^1\frac{1+(1-u)^n}{u\ln(1-u)}du \tag{50}$$

Remark: The application of formulas (48) and (50) generally follows the same guidelines as those for Poisson entropy in the previous Section 5.2.1. However, we must emphatically point out that the expressions for binomial entropy lead to an apparent ambiguity

[11] This is derived from stochastic processes with a **c.r.v.**, specifically, for the so-called gaussian *power spectral density* **PSD**. Kolmogorov showed that for a stationary Gaussian process with **PSD** $\psi_X(s)$, the differential entropy rate is given by:
$\bar{\mathbf{H}}(X) = \frac{1}{2}\log(2\pi e)+\frac{1}{4\pi}\int_0^{2\pi}\log(\psi_X(s))ds$. From what can be derived $H_{Dist}(X)\le\frac{1}{2}\log\left[2\pi e\left(\mathbf{Var}(\mathbf{X})+\frac{1}{12}\right)\right]$.

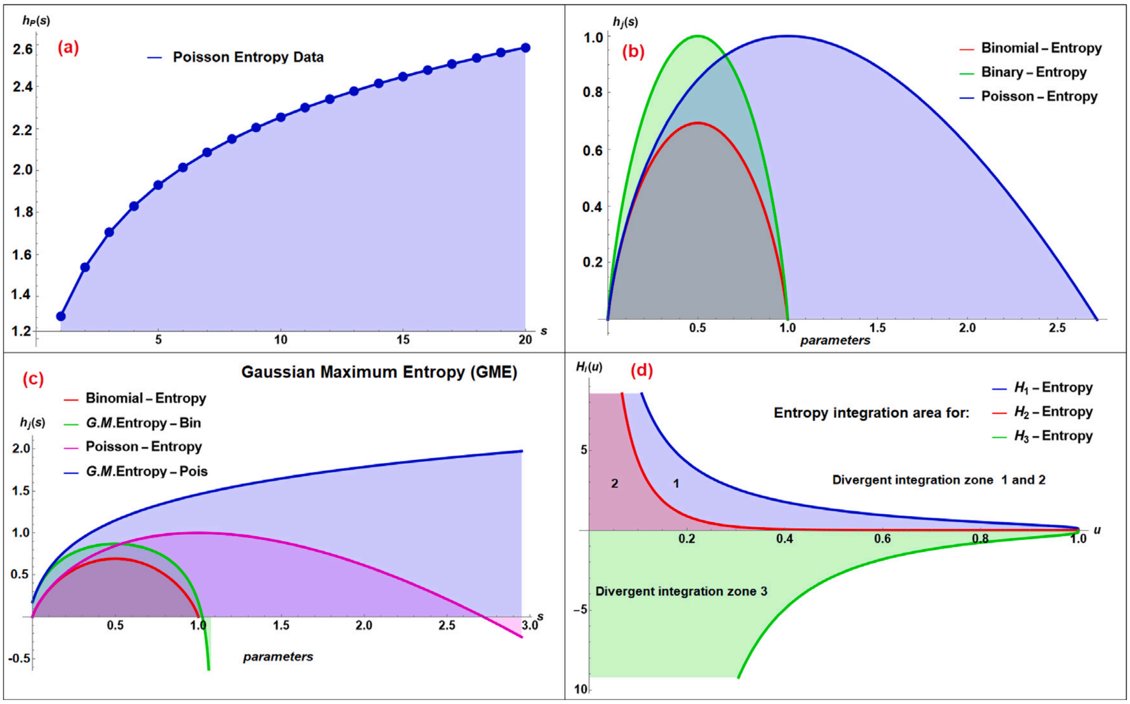

Fig. 4. Poisson and binomial entropy: In (a), we show the simulation data by numerical integration of (41), where $s = \langle s_n \rangle$ is the average value —for $s = 1, 2, \dots, 10$— and u an auxiliary variable in the integral. In (b), only the *binary type* terms independently of the terms under the integral are plotted, that is, from (41) $s - s\log(s)$, from (48) $-p\log(p) - (1-p)\log(1-p)$ in natural basis. And from (45) $h(p) = -p\log_2(p) - (1-p)\log_2(1-p)$ in base 2 —the binary base originally used in Shannon information theory—, all as a function of their parameters. (c) Here we show the same plots as in (b) —except binary entropy— superimposing the *differential entropy bounded curve* through the ***Gaussian Maximum Entropy*** **(GME)**, (which are bounds based on the normal distribution, see 5.2.1); that is, $H_{Poi}(X) \leq \frac{1}{2}\log\left[2\pi e\left(s + \frac{1}{12}\right)\right]$, and $H_{bin}(X) \leq \frac{1}{2}\log\left[2\pi e\left(s(1-s) + \frac{1}{12}\right)\right]$. Where we use: $H_{Dist}(X) \leq \frac{1}{2}\log\left[2\pi e\left(\mathbf{Var}(\mathbf{X}) + \frac{1}{12}\right)\right]$. For a more complete description about **GME**, see the references: [41], chapters 8, and 12. And [46]. Finally, in (d), we show the functions of the formula integral (48) and it also applies to (49) and (50). Numerical evaluation is not possible because they are divergent integrals (see text for full explanation, 5.2.2).

in the explicit (numerical) calculation of integrals —which we denote as $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$— since they are divergent.[12] Thus, our proposal to understand convergence (if it exists) more quickly is to use the $\mathbf{H}_j$ expressions, of which we plot the functions within the integral in Fig. 4 (**d**). This allows us to visualize the so-called areas under the curve. From this, we can deduce that the most pathetic case is area (zone) 3 (for $\mathbf{H}_3 < 0$) since its *area is negative infinite* and overwhelmingly cancels out the other two areas (zones) 1 and 2, which are much smaller.[13] Nevertheless, for numerical computation, we can resort to a trick by setting a very small fixed lower limit close to zero —asymptotically close to 0—. We can conduct *numerical experiments* that allow us to observe the overall behavior of the three involved areas and how they cancel out. From this, we can notice that zones 1 and 2 are very small, while zone three (negative) acquires increasingly large values. We could make an integration interval $[\delta, 1]$ for a δ sufficiently small, e.g.

$$\lim_{\delta\to 0}\left\{-\int_\delta^1 \frac{\left[1-u(1-e^{-\lambda t})\right]^n}{\ln(1-u)}du - \int_\delta^1 \frac{\left[1-ue^{-\lambda t}\right]^n}{\ln(1-u)}du + \int_\delta^1 \frac{1+(1-u)^n}{u\ln(1-u)}du\right\}$$

However, the third integral continues to overwhelmingly outperform the first two integrals. Fortunately, for large values of $n = N_0$, the binomial entropy (48) with the term $nh(p)$ dominates over the other three. Therefore, the dominant binomial entropy for large N_0 is the binary-like part. For radioactive substances (radioisotopes), we speak of about N_0 on the order of 10^{23}.

6. Conclusion

In this paper, we explore nuclear radioactive decay from various angles and perspectives, briefly touching upon the deterministic aspect and delving significantly into the stochastic aspect. We provide a concise analysis of decay within a deterministic framework

[12] It is worth mentioning that in the paper of M. Cheraghchi [43], the expressions for binomial entropy, formulas (17) (similar to ours (48)) and (19) are quite general. However, the numerical calculation of it is quite cumbersome. Since both formulas are infinite, but especially in formula (19), the second term is expressed as an *infinite sum of improper integrals*. We wouldn't know if that would have a rapid convergence.

[13] Another option that might be tentative is to compute the integrals as a *principal value integral* (Cauchy Integrals). Unfortunately, this only applies to areas under the curve that are symmetric or antisymmetric about the y-axis. In order to add (finite value) or cancel the infinite asymptotic zones.

in Section 2, where it is modeled as a phenomenon featuring exponential decay. However, our primary focus lies on the statistical and probabilistic aspects of decay. We commence with the concept of the so-called Radioactive Dice (a toy model; see [4–6,47]), where statistical techniques such as the Moment Generating Function (MGF) method are employed to generalize the elementary system [8,9]. These techniques serve as a foundation to advance our study using stochastic models.

In particular, we analyze the Binomial and Poisson processes within a stochastic context (Section 3). Given that radioactive decay is inherently statistical (random) in nature, we cannot predict when any of the radioisotopes will decay. To deduce the stochastic Poisson process (Section 4.3), we employ the Kolmogorov stochastic differential equations approach (in continuous time), which is well-known in the fields of statistical and quantum physics as master equations. This process is crucial for understanding systems where events of a random nature are studied, such as the counting of atomic and subatomic particles (in nuclear and atomic physics) or photons (quantum optics). By adopting this approach, we gain a better understanding of the nature of radioactive decay, considering it as a stochastic phenomenon rather than solely focusing on its counting statistics.

The literature on statistical physics, quantum optics, nuclear medicine, and related fields abounds in studies on the statistical counting of detectable particles —be they *bosons* and/or *fermions*—(see references cited in the text). However, as previously mentioned, these studies primarily focus on pure statistical counting. Specifically, they examine the statistics provided by the binomial and Poisson distributions and make comparisons between them. In our study, we take a step further by analyzing these distributions not only from an elementary statistical-probabilistic standpoint but also by applying the techniques of stochastic processes.

As previously mentioned, stochastic modeling of systems and phenomena of a random nature, such as nuclear radioactive decay, is endowed by powerful mathematical tools, such as Markov processes and Kolmogorov differential equations (master equations). These tools provide us with deeper insights into random events through the tripartite application of stochastic processes: namely, Binomial, Poisson, and exponential. These processes involve distributions of a discrete random variable (**d.r.v.**) (with the exponential being a continuous random variable) in continuous time. This enables us to comprehend how these variables evolve by conducting a comparative analysis of their behavior through the stochastic evolution of the Binomial and Poisson processes. It's important to note that they also undergo significant changes in their radioactive decay due to real physical parameters, such as decay constants and half-lives.

One of the most notable outcomes of our study was the calculation of the Poisson process modeled through the so-called pure death process, as modeled with the master equations (21), with the solution provided by (22) fully deduced in Appendix B. Introducing the sojourn time (waiting time) (Section 4.3.1 and Fig. 1), we obtain the original solution (24), which differs from those studied in the literature. Another important result is the distinction between this model and the one most commonly used in the literature, namely the Poisson process conceived as a pure birth process (Section 4.4). In Fig. 2, we present a comparison of the binomial processes proposed by Forster and Huestis (Section 3.2), as well as the Poisson processes obtained in this work.

A more fundamental approach that we further delve into is the analysis of relative fluctuations using the Fano factor (Section 5.1 and Fig. 3), as well as the variability of uncertainty through the entropy, within the context of Shannon's entropy, of Poisson and binomial distributions. We derived Eq. (41), as mentioned earlier (Section 5.2), which was previously calculated by other authors [43,44]. Nonetheless, we have calculated it based on fundamental principles using our methodologies, and we have presented the Stochastic Entropy as a function of time, namely Eq. (43). Similarly, we obtained the binomial entropy through first principles, resulting in Eqs. (48) and (50). Our approach is rooted in a theoretical framework, providing deeper insights than solely examining statistical characteristics. Entropy, a fundamental mathematical concept in thermodynamics, statistical physics, and information theory, offers deeper knowledge about random systems beyond statistical data and tests alone. Furthermore, we can explore additional metrics such as the *rate and production of entropy*, *relative entropy*, and more (which we leave for future work).

Both the detection and the associated statistics of radioactive decay entail a more complex feature when attempting to understand their evolution throughout the random physical process. As mentioned earlier, this process is divided into a deterministic collective behavior when most of the radioisotopes are in a waiting time (as determined by the *sojourn time*), without decaying ($N_0 \gg n$), and another phase when the radioisotopes decay and disintegrate at a constant yet random rate. It remains unpredictable which of all the particles will decay at any given instant of time. Therefore, the statistics and their distribution for this physical process of radioactive decay are updated in each interval of time, dividing the process into two stages: a larger, more or less deterministic portion of the radioactive substance remains without decaying during the waiting or sojourn time, while the other, smaller and more random portion undergoes progressive decay, as demonstrated in this paper. Furthermore, it should be noted that a better understanding of the radioactive decay phenomenon can be achieved through stochastic processes and statistical analysis.

CRediT authorship contribution statement

Sergio Sánchez-Sánchez: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Ernesto Cortés-Pérez:** Validation, Software, Investigation, Data curation. **Víctor I. Moreno-Oliva:** Software, Resources, Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Appendix A

In this appendix we carry out the demonstrations of Eqs. (27) and (28) of Section 4.4. We start with the case of two independent Poisson random variables, $X_j \sim Poisson(X_j, \lambda_j) = P_{X_j}(x_j)$, $(j = 1, 2)$ with different parameters. Then we proceed to generalize for n r. v. of Poisson, i.e. $X_j \sim Poisson(X_j, \lambda_j) = P_{X_j}(x_j)$, $(j = 1, 2, \ldots, n)$. All this using the discrete convolution formula for 2 and n probability distribution functions.

$$
\begin{aligned}
P_{S_2}(s_2) &= P_{X_1+X_2}(s_2) = P_{X_1}(x_1) * P_{X_2}(x_2) \\
&= \sum_{x_1,x_2}^{s_2} \frac{e^{-\lambda_1}\lambda^{x_1}}{(x_1)!}\frac{e^{-\lambda_2}\lambda^{x_2}}{(x_2)!} = \sum_{x_1=0}^{s_2} \frac{e^{-\lambda_1}e^{-\lambda_2}}{(x_1)!}\frac{\lambda^{x_1}\lambda^{s_2-x_1}}{(s_2-x_1)!} \\
&= \sum_{x_1=0}^{s_2} e^{-(\lambda_1+\lambda_2)}\frac{[\lambda^{x_1}\lambda^{s_2-x_1}]}{(x_1)!(s_2-x_1)!}\underbrace{\left(\frac{(s_2)!}{(s_2)!}\right)}_{1} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{(s_2)!}\sum_{x_1=0}^{s_2}\underbrace{\frac{(s_2)!}{(x_1)!(s_2-x_1)!}}_{\text{binomial coefficient}}\left[\lambda^{x_1}\lambda^{s_2-x_1}\right] = \frac{e^{-(\lambda_1+\lambda_2)}}{(s_2)!}\sum_{x_1=0}^{s_2}\binom{s_2}{x_1}\lambda^{x_1}\lambda^{s_2-x_1} \\
P_{S_2}(s_2) &= \frac{e^{-(\lambda_1+\lambda_2)}}{(s_2)!}\left(\lambda_1+\lambda_2\right)^{s_2}
\end{aligned}
\tag{51}
$$

Now we proceed to show how Eq. (28) is obtained, generating for n independent Poisson r. v.

$$
\begin{aligned}
P_{S_n}(s_n) &= P_{X_1+X_2+\cdots+X_n}(s_n) = P_{X_1}(x_1) * P_{X_2}(x_2) * \cdots * P_{X_n}(x_n) \\
&= \sum_{\substack{x_1,\ldots,x_n \\ s_n=x_1+x_2\cdots+x_n}}^{s_n} P_{X_1}(x_1)P_{X_2}(x_2)\cdots P_{X_n}(x_n) \\
&= \sum_{\substack{x_1,\ldots,x_n \\ s_n=x_1+x_2\cdots+x_n}}^{s_n} \frac{\lambda_1^{x_1}e^{-\lambda_1}}{(x_1!)}\frac{\lambda_2^{x_2}e^{-\lambda_2}}{(x_2!)}\cdots\frac{\lambda_n^{x_n}e^{-\lambda_n}}{(x_n!)} \\
&= \sum_{\substack{x_1,\ldots,x_n \\ s_n=x_1+x_2\cdots+x_n}}^{s_n} \frac{e^{-\lambda_1}e^{-\lambda_2}\cdots e^{-\lambda_{n-1}}e^{-\lambda_n}}{x_1!x_2!\cdots x_{n-1}!x_n!}\left[\lambda_1^{x_1}\lambda_2^{x_2}\cdots\lambda_{n-1}^{x_{n-1}}\lambda_n^{x_n}\right] \\
&= e^{-(\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_n)}\sum_{\substack{x_1,\ldots,x_n \\ s_n=x_1+x_2\cdots+x_n}}^{s_n}\left\{\frac{\lambda_1^{x_1}\lambda_2^{x_2}\lambda_3^{x_3}\cdots\lambda_n^{x_n}}{x_1!x_2!x_3!\cdots x_n!}\right\}\underbrace{\left(\frac{(s_n)!}{(s_n)!}\right)}_{1} \\
P_{S_n}(s_n) &= \frac{e^{-(\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_n)}}{s_n!}\sum_{\substack{x_1,\ldots,x_n \\ s_n=x_1+x_2\cdots+x_n}}^{s_n}\underbrace{\frac{s_n!}{x_1!x_2!x_3!\cdots x_n!}}_{\text{multinomial coefficient}}\left(\lambda_1^{x_1}\lambda_2^{x_2}\lambda_3^{x_3}\cdots\lambda_n^{x_n}\right) \\
&= \frac{e^{-(\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_n)}}{s_n!}\sum_{\substack{x_1,\ldots,x_n \\ s_n=x_1+x_2\cdots+x_n}}^{s_n}\binom{s_n}{x_1,x_2,\ldots,x_n}\left(\lambda_1^{x_1}\lambda_2^{x_2}\lambda_3^{x_3}\cdots\lambda_n^{x_n}\right) \\
&= \frac{e^{-(\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_n)}}{s_n!}\left(\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_n\right)^{s_n} \\
P_{S_n}(s_n) &= \frac{\exp\left[-\sum_{j=1}^{n}\lambda_j\right]}{s_n!}\left(\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_n\right)^{s_n}
\end{aligned}
\tag{52}
$$

This shows Eq. (28), which, already we said above, is the *probability mass function (pmf)* for n random variables distributed by Poisson Distribution, with parameters $\lambda_j = \langle x_j \rangle$ *the mean (expected value)*. Notice that $s_n!$ represents the product of all random variables $x_j!$, that is, $s_n! = x_1!x_2!\cdots x_n!$.

Appendix B

Here we carry out the solution of Eqs. (22) —it would be similar for (24)—. Solving the first equation of (21), we obtain

$$P_{N_0}(t) = e^{-\mu t}$$

Then iteratively solving the other two equations of (21), for $n = 1, 2, \dots, N_0 - 1$. Using the *integrating factor* technique, we get (with integrating factor: $h(t) = \mu t$)

$$\begin{aligned}
P_n(t) &= e^{-h(t)}\left\{\int_0^t \mu e^{h(s)} P_{n+1}(s)ds + c\right\}; \quad (c = 0)\\
P_n(t) &= e^{-\mu t}\int_0^t \mu e^{\mu s} P_{n+1}(s)ds\\
P_n(t) &= e^{-\mu t} R_n(t)\\
&\text{where}\\
R_n(t) &= \int_0^t \mu e^{\mu s} P_{n+1}(s)ds, \quad \Longrightarrow R_n(t) = e^{\mu t} P_n(t)\\
\Longrightarrow \quad \frac{dR_n(t)}{dt} &= R'_n(t) = \mu R_{n-1}(t)
\end{aligned} \tag{53}$$

Solving a few terms of formula (41) with the help of (40), we have

$$\begin{aligned}
R'_{N_0-1}(t) &= \mu R_{N_0-2}(t) \quad \Longrightarrow \quad R_{N_0-1}(t) = \mu t = \frac{\mu t}{1!}\\
R'_{N_0-2}(t) &= \mu R_{N_0-3}(t) \quad \Longrightarrow \quad R_{N_0-2}(t) = \frac{(\mu t)^2}{1\cdot 2} = \frac{(\mu t)^2}{2!}\\
R'_{N_0-3}(t) &= \mu R_{N_0-4}(t) \quad \Longrightarrow \quad R_{N_0-3}(t) = \frac{(\mu t)^3}{1\cdot 2\cdot 3} = \frac{(\mu t)^3}{3!}\\
&\vdots\\
R'_{N_0-k}(t) &= \mu R_{N_0-k+1}(t) \quad \Longrightarrow \quad R_{N_0-k}(t) = \frac{(\mu t)^k}{1\cdot 2\cdot 3\cdots(k-1)\cdot k} = \frac{(\mu t)^k}{k!}\\
&\vdots\\
R'_{N_0-n}(t) &= \mu R_{N_0-n+1}(t) \Longrightarrow \quad R_{N_0-n}(t) = \frac{(\mu t)^{N_0-n}}{1\cdot 2\cdots(N_0-n-1)\cdot(N_0-n)} = \frac{(\mu t)^{N_0-n}}{(N_0-n)!}
\end{aligned}$$

Then we can get

$$P_n(t) = \frac{(\mu t)^{N_0-n}}{(N_0-n)!}e^{-\mu t}, \quad n = 1, 2, \dots, k, \dots, N_0 \tag{54}$$

On the other hand, it must be fulfilled (Kolmogorov's axiom), for each t we have

$$\sum_{n=0}^{N_0} P_n(t) = 1, \quad \Longrightarrow \quad P_0(t) + \sum_{n=1}^{N_0} P_n(t) = 1$$

we get

$$P_0(t) = 1 - \sum_{j=1}^{N_0}\frac{(\mu t)^{N_0-j}}{(N_0-j)!}e^{-\mu t} = 1 - \sum_{n=0}^{N_0-1}\frac{(\mu t)^n}{n!}e^{-\mu t} \tag{55}$$

where we made the change: $n = N_0 - j \quad \Rightarrow j = N_0 - n = 1$ in the last line, with n running from $n = 0, 1, 2, \dots, N_0 - 1$. With this, we prove the solutions for Eqs. (22) and (24).

Appendix C

Poisson stochastic entropy function

In this section, we develop more details of the explicit calculation of the stochastic entropy functions of Poisson and Binomial.

$$\begin{aligned}
\mathbf{H}(X_{Poisson}) &= \mathbf{H}_{Poiss} = -\sum_{k=0}^{\infty}\left(\frac{s^k}{k!}e^{-s}\right)\log\left(\frac{s^k}{k!}e^{-s}\right)\\
&= -\sum_{k=0}^{\infty}\frac{s^k}{k!}e^{-s}\Big[k\log(s) - \log(k!) - s\Big]\\
&= -e^{-s}\left\{\sum_{k=0}^{\infty}\frac{s^k}{k!}k\log(s)\right\} + se^{-s}\underbrace{\left[\sum_{k=0}^{\infty}\frac{s^k}{k!}\right]}_{e^s} + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\big[\log(k!)\big]\\
&= -e^{-s}\log(s)\underbrace{\left\{\sum_{k=0}^{\infty}\frac{s^k}{k!}k\right\}}_{\text{see note:}\star} + se^{-s}e^{s} + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\big[\log(k!)\big]\\
&= -e^{-s}\log(s)\big[e^s + se^s - e^s\big] + s + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\big[\log(k!)\big]
\end{aligned}$$

$$\mathbf{H}_{Poiss} = \left(s - s\log(s)\right) + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\left[\log\Gamma(k+1)\right] \tag{56}$$

From this, we can obtain Eqs. (41) and (43).

Math note: ($\star$) We can split this series into two parts, as indicated in the following development (Using formula **1.212**, page 26, from Ref. [45]):

$$e^s(1+s) = \sum_{k=0}^{\infty}\frac{s^k(k+1)}{k!} = \sum_{k=0}^{\infty}\frac{s^k}{k!}k + \sum_{k=0}^{\infty}\frac{s^k}{k!} = \sum_{k=0}^{\infty}\frac{s^k}{k!}k + e^s \quad\Longrightarrow\quad \sum_{k=0}^{\infty}\frac{s^k}{k!}k = e^s + se^s - e^s = se^s$$

Now we show how to transform the improper (semi-infinite) integral to the definite integral on the interval $[0,1]$ in (42).

$$\log\Gamma(k+1) = \int_0^{\infty}\left[k - \frac{1-e^{-kx}}{1-e^{-x}}\right]\frac{e^{-x}}{x}dx$$

We make the change of variable $u = e^{-x} \Rightarrow x = -\ln u$, or $1/x = -1/\ln u$. Thus $du = -e^{-x}dx = -udx$, $\Rightarrow dx = -\frac{du}{u}$. Also, we have $u(0) = e^{-(0)} = 1$, and $u(\infty) = e^{-(\infty)} = 0$, [for this case $\log(\cdot) = \ln(\cdot)$]. Substituting into the previous improper integral (i.e. from (42)), we obtain the definite integral:

$$\int_{u(0)=1}^{u(\infty)=0}\left[k - \frac{1-u^k}{1-u}\right]\frac{u}{(-\ln u)}\left(-\frac{du}{u}\right) = -\int_0^1\left[k - \frac{1-u^k}{1-u}\right]\frac{du}{\ln(u)}$$

$$\Longrightarrow\quad \log\Gamma(k+1) = \ln\Gamma(k+1) = \int_0^1\left[\frac{1-u^k}{1-u} - k\right]\frac{du}{\ln(u)}$$

Substituting into Eq. (56) —that is, initially into (41)—

$$\begin{aligned}
\mathbf{H}_{Poiss} &= \left(s - s\log(s)\right) + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\left[\log\Gamma(k+1)\right]\\
&= \left(s - s\log(s)\right) + e^{-s}\sum_{k=0}^{\infty}\frac{s^k}{k!}\int_0^1\left[\frac{1-u^k}{1-u} - k\right]\frac{du}{\ln(u)}\\
&= \left(s - s\log(s)\right) + e^{-s}\int_0^1\sum_{k=0}^{\infty}\frac{s^k}{k!}\left[\frac{1-u^k}{1-u} - k\right]\frac{du}{\ln(u)}\\
&= \left(s - s\log(s)\right) + e^{-s}\int_0^1\left\{\left[\frac{1}{1-u}\sum_{k=0}^{\infty}\frac{s^k}{k!}(1-u^k)\right] - \left(\sum_{k=0}^{\infty}\frac{s^k k}{k!}\right)\right\}\frac{du}{\ln(u)}\\
&= \left(s - s\log(s)\right) + \int_0^1\left\{\frac{e^{-s}}{1-u}\left[\sum_{k=0}^{\infty}\frac{s^k}{k!} - \sum_{k=0}^{\infty}\frac{(su)^k}{k!}\right] - e^{-s}\underbrace{\left(\sum_{k=0}^{\infty}\frac{s^k k}{k!}\right)}_{se^s}\right\}\frac{du}{\ln(u)}\\
&= \left(s - s\log(s)\right) + \int_0^1\left\{\frac{e^{-s}}{1-u}\left[e^s - e^{su}\right] - e^{-s}\left(se^s\right)\right\}\frac{du}{\ln(u)}\\
&= \left(s - s\log(s)\right) + \int_0^1\left\{\frac{1-e^{-s(1-u)}}{1-u} - s\right\}\frac{du}{\ln(u)}
\end{aligned} \tag{57}$$

With this, Eq. (41) is proved —from which we can get directly (43)—.

Appendix D

Binomial stochastic entropy function

We now show the computational details of the binomial stochastic entropy function. The standard binomial distribution is $P_n(k) = f_{Bin}(k) = B(k;n,p) = \binom{n}{k}p^k(1-p)^{n-k}$. This is a p.d.f., which has a factorial coefficient. And we can express it in terms of the Gamma function as follows

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)} \tag{58}$$

In the factor containing the logarithm in (44), we simply systematically apply the elementary logarithmic properties to it.

$$\begin{aligned}
\mathbf{H}_{Bin} = &-\underbrace{\sum_{k=0}^{\infty}\frac{\Gamma(n+1)p^k\left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)}}_{\mathbf{S}(bin)}\\
&\times\left[\log\Gamma(n+1) - \log\Gamma(k+1) - \log\Gamma(n-k+1) - k\log p + (n-k)\log\left[1-p\right]\right]
\end{aligned}$$

$$= -\mathbf{S}(bin) \times \left[\log \Gamma(n+1) - \log \Gamma(k+1) - \log \Gamma(n-k+1)\right]$$
$$- \mathbf{S}(bin) \times \left[k \log p + (n-k) \log \left[1-p\right]\right] \tag{59}$$

Then

$$-\mathbf{S}(bin)\left[\log \Gamma(n+1) - \log \Gamma(k+1) - \log \Gamma(n-k+1)\right] = -\Gamma(n+1) \log \Gamma(n+1) \sum_{k=0}^{\infty} \frac{p^k \left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)}$$
$$+ \sum_{k=0}^{\infty} \frac{\Gamma(n+1) \log \Gamma(n-k+1) p^k \left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)} + \sum_{k=0}^{\infty} \frac{\Gamma(n+1) \log \Gamma(k+1) p^k \left(1-p\right)^{n-k}}{\Gamma(k+1)\Gamma(n-k+1)}$$

So another term

$$-\mathbf{S}(bin)\left[k \log p + (n-k) \log \left[1-p\right]\right] = -\log p \sum_{k=0}^{\infty} k \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)} p^k \left(1-p\right)^{n-k}$$
$$- \log (1-p) \sum_{k=0}^{\infty} (n-k) \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)} p^k \left(1-p\right)^{n-k}$$
$$= -\log p \sum_{k=0}^{\infty} k \frac{n!}{k!(n-k)!} p^k \left(1-p\right)^{n-k} - \log (1-p) \sum_{k=0}^{\infty} (n-k) \frac{n!}{k!(n-k)!} p^k \left(1-p\right)^{n-k}$$

After carrying out some algebraic manipulations on the quotient factorial, we get

$$= -np \log p \underbrace{\sum_{j=0}^{\infty} \binom{n}{j} p^j \left(1-p\right)^{n-j}}_{\sum_{j=0}^{\infty} f_{Bin}(j)=1} - n(1-p) \log (1-p) \underbrace{\sum_{j=0}^{\infty} \binom{n}{j} p^j \left(1-p\right)^{n-j}}_{=1}$$

So, we get the *binary entropy* (see Eq. (45))

$$nh(p) = -np \log p - n(1-p) \log (1-p) = -n\left[p \log p + (1-p) \log (1-p)\right] \tag{60}$$

We define the expected values (46) and (47) to simplify Eq. (48), and now we show the more detailed calculation.

$$\underbrace{\mathbf{E}\left[\log \Gamma(k+1)\right]}_{E_{Bin}(n,p)} = \sum_{k=0}^{\infty} \log \Gamma(k+1) \frac{n!}{k!(n-k)!} p^k \left(1-p\right)^{n-k} \tag{61}$$

Similarly

$$\underbrace{\mathbf{E}\left[\log \Gamma(n-k+1)\right]}_{E_{Bin}(n,1-p)} = \sum_{k=0}^{\infty} \log \Gamma(n-k+1) \frac{n!}{k!(n-k)!} p^k \left(1-p\right)^{n-k} \tag{62}$$

These expected values are calculated using the following formulas developed by Cheraghchi (specially *Theorem 1*) [43]. We obtain a more general formula (see Eqs. (49) and (50)) in terms of integral functions similar to the Poisson entropy function. Next, we put the formula (2) of the Ref. [43], which we use to calculate the expected values:

$$\mathbf{E}\left[\log \Gamma(X+\alpha)\right] = \log \Gamma(\alpha) + \int_0^{\infty} \left[\frac{\mu e^{-t}}{t} - \frac{e^{-\alpha t}(1-\mathbf{M}(-t))}{t(1-e^{-t})}\right] dt \tag{63}$$

where $\mathbf{M}(t)$[14] is the moment generating function (**MGF**). In elementary probability theory (see the following references: [9,18,27, 28,35]) the binomial distribution function has the **MGF** following

$$\mathbf{M}_X(t) = \sum_{m=0}^{n} e^{tm} \binom{n}{m} p^m \left(1-p\right)^{n-m} = \left[pe^t + q\right]^n = \left[pe^t + (1-p)\right]^n \tag{64}$$

Thus for $\mathbf{M}_X(-t) = \left[pe^{-t} + (1-p)\right]^n$. With $\mu = np$, $\alpha = 1 \Rightarrow \log \Gamma(\alpha = 1) = \log(1) = 0$, so

$$E_{Bin}(n,p) = \int_0^{\infty} \left[\frac{npe^{-t}}{t} - \frac{e^{-t}\left[1 - \left(pe^{-t} + (1-p)\right)^n\right]}{t(1-e^{-t})}\right] dt \tag{65}$$

Similarly for

$$E_{Bin}(n,1-p) = \int_0^{\infty} \left[\frac{ne^{-t}}{t} - \frac{npe^{-t}}{t} - \frac{e^{-t}\left[1 - \left(p - (1-p)e^{-t}\right)^n\right]}{t(1-e^{-t})}\right] dt \tag{66}$$

[14] We must point out that in this case, t in the **MGF** is only a parameter (dummy variable), and does not represent time. This comes from elementary probability theory (see references, and also see Section 2.1)

Now we proceed to add both expressions (64) and (66), and we subtract the term of the Gamma function (42); that is, $-\log\Gamma(n+1)+E_{Bin}(n,p)+E_{Bin}(n,1-p)$ as in Eq. (48). And after carrying out several simplifications we obtain

$$\begin{aligned}
&-\log\Gamma(n+1)+\Big[E_{Bin}(n,p)+E_{Bin}(n,1-p)\Big]=-\int_0^\infty\left[n-\frac{1-e^{-nt}}{1-e^{-t}}\right]\frac{e^{-t}}{t}dt\\
&+\int_0^\infty\left[-\frac{e^{-t}\left[1-\left(p-(1-p)e^{-t}\right)^n\right]}{t(1-e^{-t})}+\frac{ne^{-t}}{t}-\frac{e^{-t}\left[1-\left(p-(1-p)e^{-t}\right)^n\right]}{t(1-e^{-t})}\right]dt\\
&=\int_0^\infty\left[\frac{\left[e^{-t}\left(1-p+pe^{-t}\right)^n\right]}{t(1-e^{-t})}+\frac{e^{-t}\left[p+(1-p)e^{-t}\right]^n}{t(1-e^{-t})}-\frac{e^{-t}}{t(1-e^{-t})}-\frac{e^{-nt}e^{-t}}{t(1-e^{-t})}\right]dt\\
&=\int_0^\infty\left[\frac{\left[e^{-t}\left(1-p+pe^{-t}\right)^n\right]}{e^{-t}t(e^t-1)}+\frac{e^{-t}\left[p+(1-p)e^{-t}\right]^n}{e^{-t}t(e^t-1)}-\frac{e^{-t}}{e^{-t}t(e^t-1)}-\frac{e^{-t}e^{-nt}}{e^{-t}t(e^t-1)}\right]dt\\
&=\underbrace{\int_0^\infty\left[\frac{\left(1-p+pe^{-t}\right)^n}{t(e^t-1)}\right]dt}_{H_1}+\underbrace{\int_0^\infty\left[\frac{\left[p+(1-p)e^{-t}\right]^n}{t(e^t-1)}\right]dt}_{H_2}-\underbrace{\int_0^\infty\left[\frac{(1+e^{-nt})}{t(e^t-1)}\right]dt}_{H_3}
\end{aligned}\tag{67}$$

So, we have

$$\mathbf{H}_{Bin}(X)=nh(p)+\mathbf{H}_1+\mathbf{H}_2-\mathbf{H}_3\tag{68}$$

We show the development and transformation —of the integral semi-infinite to the interval $[0,1]$— of the *partial entropies* $\mathbf{H}_1$ ($\mathbf{H}_2$, similarly) and $\mathbf{H}_3$. Transforming the integrand once again as: $\left[1-p(1-e^{-t})\right]$ and $\left(e^t-1\right)=e^t\left(1-e^{-t}\right)$.

$$\mathbf{H}_1=\int_0^\infty\left[\frac{\left(1-p+pe^{-t}\right)^n}{t(e^t-1)}\right]dt=\int_0^\infty\frac{\left[1-p(1-e^{-t})\right]^n}{te^t\left(1-e^{-t}\right)}dt$$

Once again, we make the following variable changes, to transform the integral: $u=1-e^{-t}\Rightarrow du=e^{-t}dt$. But we also have to $e^{-t}=1-u$, and $e^t=\frac{1}{1-u}$, hence $t=\ln\left[\frac{1}{1-u}\right]=-\ln\left(1-u\right)\Rightarrow\quad dt=\frac{du}{e^{-t}}=\frac{du}{(1-u)}$. For the integral limits we have: $u(0)=1-e^{-(0)}=0$, and $u(\infty)=1-e^{-(\infty)}=1-0=1$, and plugging it into the integral,

$$\mathbf{H}_1=\int_0^\infty\frac{\left[1-p(1-e^{-t})\right]^n}{te^t\left(1-e^{-t}\right)}dt=\int_0^1\frac{\left[1-pu\right]^n}{-\ln\left(1-u\right)}\left(1-u\right)\frac{du}{\left(1-u\right)}=-\int_0^1\frac{\left[1-pu\right]^n}{\ln\left(1-u\right)}du\tag{69}$$

Similarly, we can transform $\mathbf{H}_2$:

$$\mathbf{H}_2=\int_0^\infty\left[\frac{\left[p+(1-p)e^{-t}\right]^n}{t(e^t-1)}\right]dt=-\int_0^1\frac{\left[1-(1-p)u\right]^n}{\ln\left(1-u\right)}du\tag{70}$$

And for $\mathbf{H}_3$:

$$\mathbf{H}_3=\int_0^\infty\left[\frac{(1+e^{-nt})}{t(e^t-1)}\right]dt=\int_0^1\frac{1+\left(1-u\right)^n}{u\ln\left(1-u\right)}du\tag{71}$$

Finally, we get the full binomial entropy functions (48) —from this Eqs. (49) and (50)—,

$$\mathbf{H}_{Bin}(X)=nh(p)-\int_0^1\frac{\left[1-pu\right]^n}{\ln\left(1-u\right)}du-\int_0^1\frac{\left[1-(1-p)u\right]^n}{\ln\left(1-u\right)}du+\int_0^1\frac{1+\left(1-u\right)^n}{u\ln\left(1-u\right)}du\tag{72}$$

References

[1] R.C. Hannum, A.N. Cabot, Practical Casino Math, second ed., in: Casino Management Series, Trace Publication, 2005, https://ajp.aapt.org/.

[2] L. Mandel, E. Wolf, Optical Coherence and Quantum Optics, first ed., Cambridge University Press, 1995.

[3] R. Mahnke, J. Kaupuzs, I. Lubashevsky, Physics of Stochastic Processes: How Randomness Acts in Time, WILEY-VCH Verlag GmbH & Co., 2009.

[4] Murray Arthur, Hart Ian, The 'radioactive dice' experiment: why is the 'half-life' slightly wrong? in: Physics Education, Vol. 47, IOP Publishing, 2012, pp. 197–201, (2).

[5] Sarmistha Sahu, Radioactive branching using dice, Creat. Educ. 3 (5) (2012) 671–673, http://dx.doi.org/10.4236/ce.2012.35099.

[6] Emeric Schultz, Dice shaking as an analogy for radioactive decay and first order kinetics, J. Chem. Educ. 74 (5) (1997) 505, http://dx.doi.org/10.1021/ed074p505.

[7] S.P. Huestis, Understanding the origin and meaning of the radioactive decay equation, J. Geosci. Educ. 50 (5) (2002) 505, http://dx.doi.org/10.5408/1089-9995-50.5.524.

[8] A.K. Singh, R.J. Dalpatadu, A.F. Lucas, The probability distribution of the sum of several dice: Slot applications, UNLV Gaming Res. Rev. J. (2011) https://digitalscholarship.unlv.edu/grrj/vol15/iss2/10.

[9] S. Sánchez-Sánchez, E. Péres-Cortes, V.I. Moreno-Oliva, B.D. Pacheco, A.F. Aguilar, Quick counting of probability distributions for the sum of multi-sided dice: a physical and didactic approach, Eur. J. Phys. (2022) http://dx.doi.org/10.1088/1361-6404/ac6913.

[10] W.S.C. Williams, Nuclear and Particle Physics, Oxford University Press, Oxford, 1991.

[11] Kenneth S. Krane, Introductory Nuclear Physics, John Wiley and Sons, Inc., 1988.

[12] A. Sitek, A.M. Celler, Limitations of Poisson statistics in describing radioactive decay, Eur. J. Med. Phys. 31 (2015) 11051107.

[13] Sitek Arkadius, Statistical Computing in Nuclear Imaging, CRC Press Taylor & Francis Group, 2015.

[14] J. Foster, K. Kouris, I.P. Matthews, N.M. Spyrou, Binomial vs. Poisson statistics in radiation studies, Nucl. Instrum. Methods Phys. Res. 212 (1) (1983) 301–305.
[15] N.M. Spyrou, J. Foster, M.C. Jones, K. Kouris, I.P. Matthews, Should the Poisson statistical density function be used in the mesurement of short-lived isotopes? J. Radio-Anal. Chem. 61 (1) (1981) 121–130.
[16] J. Wang, P. Shang, W. Ge, Multifractal cross-correlation analysis based on statistical moments, Fractals: Complex Geom. Patterns Scaling Nat. Soc. 20 (03n04) (2012) 271–279.
[17] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H. Stanley, Multifractal detrended fluctuation analysis of non-stationary time series, Phys. A 316 (1) (2022) 87–114.
[18] H. Kobayashi, B.L. Mark, W. Turin, Probability, Random Processes, and Statistical Analysis, first ed., Cambridge University Press, 2012.
[19] R. Chandra, A. Rahmim, Nuclear Medicine Physics: The Basics, eighth ed., Ed. Wolters-Kluwer, Philadelphia, 2018.
[20] D.L. Bailey, J.L. Humm, A. Todd-Pokropek, A.Van. Aswegen, Nuclear Medicine Physics: A Handbook for Students and Teachers, International Atomic Energy Agency, Vienna, 2014.
[21] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (4) (1948) 623–656.
[22] W.R. Leo, Techniques for Nuclear and Particle Physics Experiments: An how-to Approach, first ed., Springer-Verlag, 1987.
[23] G.F. Knoll, Radiation Detection and Measurement, fourth ed., in: Nuclear Engineering Series, John Wiley and Sons, Inc, 2010.
[24] A. Malyzhenkov, V. Lebedev, A. Castro, Nuclear recoil spectroscopy of levitated particles, Phys. Rev. A 98 (2018) 052103.
[25] P.T. Greenland, Seeking non-exponential decay, Nature 335 (6188) (1988) 298.
[26] E.B. Norman, S.B. Gazes, S.G. Crane, D.A. Bennett, Tests of the exponential decay law at short and long times, Phys. Rev. Lett. 60 (1988) 2246–2249.
[27] K.M. Ramachandran, C.P. Tsokos, Mathematical Statistics with Applications, first ed., Elsevier Academic Press, 2009.
[28] D.D.D.D. Wackerly, W. Mendenhall III, R.L. Scheaffer, Mathematical Statistics with Applications, seventh ed., Thomson Brooks/Cole, 2008.
[29] I. Eliazar, Universal Poisson-process limits for general random walks, Phys. A 512 (2018) 1160–1174.
[30] P.G. Hoel, S.C. Sidney C. Port, C.J. Stone, Introduction to Stochastic Process, Hougthan Mifflin Company, 1997.
[31] J.A. Gubner, Probability and Random Process for Electrical and Computer Engineers, Cambridge University Press, 2006.
[32] S.M. Ross, Introduction to Probability Models, tenth ed., USA Academic Press is an imprint of Elsevier, Burlington, MA, 2010.
[33] H.M. Taylor, S. Karlin, An Introduction to Stochastic Modeling, third ed., Academic Press is an imprint of Elsevier, San Diego California, USA, 1998.
[34] H.M. Moya-Cessa, F. Soto-Eguibar, Differential Equations: An Operational Approach, Rinton Press, Inc., New Jersey 07652, USA, 2011.
[35] G.C. Canavos, Applied Probability and Statistical Methods, first ed., Little, Brown and Company, 1984.
[36] E. Parzen, Modern Probability Theory and its Applications, Wiley-Interscience, New York, 1992.
[37] M.H. DeGroot, M.J. Schervish, Probability and Statistics, second ed., Addison-Wesley, Reading, MA, 1986.
[38] U. Fano, Ionization yield of radiations. II. The fluctuations of the number of ions, Phys. Rev. 72 (1) (1947) 26–29.
[39] K. Rajdl, P. Lansky, L. Lubomir Kostal, Fano factor: A potentially useful information, Front. Comput. Neurosci. 14 (2020) 569049, http://dx.doi.org/10.3389/fncom.2020.569049.
[40] J.W. Shuai, S. Zeng, P. Jung, Coherence resonance: On the use and abuse of the Fano factor, Fluct. Noise Lett. 2 (3) (2002).
[41] T.M. Cover, J.A. Thomas, Elements of Information Theory, second ed., Wiley, New York, 2006.
[42] L. Landau, E.M. Lifshitz, Statistical Physics, Pergamon Press, Oxford, 1980.
[43] M. Cheraghchi, Expressions for the entropy of binomial-type distributions, in: IEEE International Symposium on Information Theory, ISIT, 2018.
[44] A. Martinez, Spectral efficiency of optical direct detection, J. Opt. Soc. Am. B 24 (4) (2007) 739–749.
[45] I.S. Gradshteyn, I.M. Ryzhik, A. Jeffrey, D. Zwillinger (Eds.), Tables of Integrals, Series, and Products, seventh ed., Academic Press, 2007.
[46] R. Tamir, Entropy rate bounds via second-order statistics, 2022, http://dx.doi.org/10.48550/arXiv.2203.05237.
[47] Chi-K. Chow, T.D. Cohen, Quantum coins, dice, and children: Probability and quantum statistics, Am. J. Phys. 68 (2000) 829.