CRM Master Research Projects

CENTRE DE RECERCA MATEMÀTICA

Probability Distribution of the Radionuclide Half-Lives

Author: Francesc Font Martínez

Advisor: Álvaro Corral Cano

Tutor: Ramon Nonell Torrent

MASTER'S DEGREE THESIS

MSc in Mathematics Engineering

Title: Probability distribution of the radionuclide half-lives.

Author: Francesc Font Martinez.

Advisor: Álvaro Corral Cano.

Institution: Centre de Recerca Matemàtica.

Tutor: Ramon Nonell Torrent.

Department: Estadística i Investigació Operativa.

Academic year: 2009/2010.

Facultat de Matemàtiques i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Probability distribution of the radionuclide half-lives

Francesc Font Martinez*

Supervised by

Álvaro Corral Cano† & Ramon Nonell Torrent‡

January 29, 2010

*Student of the MSc in Mathematics Engineering in the Facultat de Matemàtiques i Estadística from the UPC.

†Centre de Rercerca Matemàtica (Campus UAB).

‡Departament d'Estadística i Investigació Operativa from the UPC.

Contents

1 Motivation

Power law distributions, a well-known model in the theory of real random variables, characterize a wide variety of natural and man made phenomena. The intensity of earthquakes, the word frequencies, the solar flares and the sizes of power outages are distributed according to a power law distribution.

Recently, given the usage of power laws in the scientific community, several articles have been published criticizing the statistical methods used to estimate the power law behaviour and establishing new techniques to their estimation with proven reliability.

The main object of the present study is to go in deep understanding of this kind of distribution and its analysis, and introduce the half-lives of the radioactive isotopes as a new candidate in the nature following a power law distribution, as well as a "canonical laboratory" to test statistical methods appropriate for long-tailed distributions.

2 Introduction

Power laws distributions have appeared in a large variety of research fields as ecology, physics, economy and finance, etc—see [1], [2], [3]. But sometimes the methods used for the power law distribution analysis in the scientific literature have not been as accurate as should be—see [1], [4], [2].

A power law density distribution is a function of the form $f(x) = Cx^{-\alpha}$, $C > 0$ and being the exponent a positive constant parameter. Normally, in empirical cases, the value of the exponent is unknown and it has to be estimated. A mechanism to estimate this value, for a supposed quantity x obeying a power law distribution, is plotting the logarithm of the empirical density versus the logarithm of the quantity and estimate the slope of the plot by a linear regression. Given that, $\ln f(x) = -\alpha \ln x + \text{constant}$.

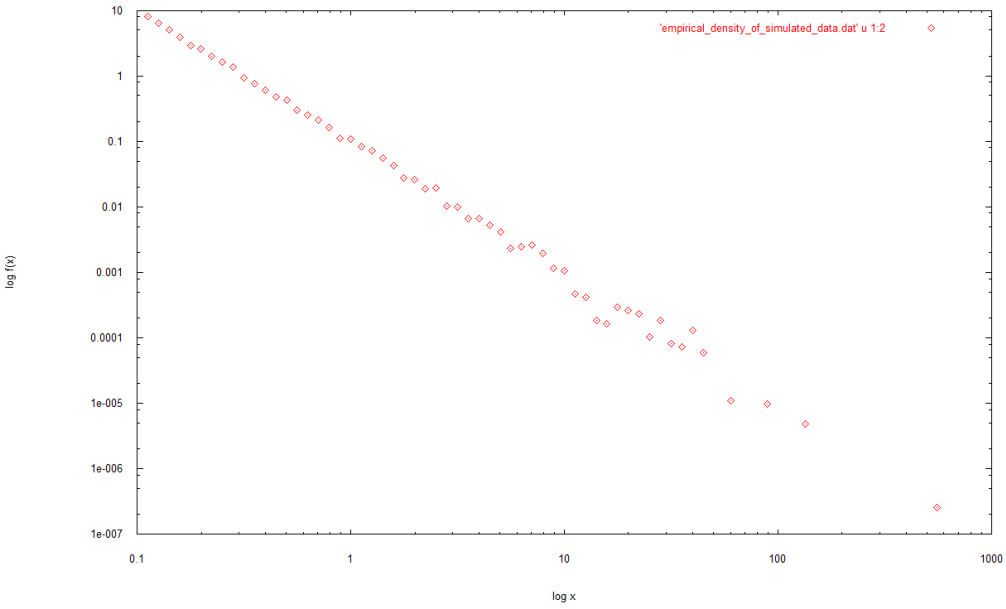


Figure 1: Empirical probability density of a simulated power law with exponent $\alpha = 2.0$, $x_{\min}$, and 10000 random data points, represented in a double logarithmic scale. The density is estimated following the recipe of [5].

Recent studies have proved that this method generates significant inaccuracy in the estimation of the exponent of the power law—see [4]. Therefore, as would be usual in the estimation of parameters of probability distributions, the estimation of α should be done with a robust method as the maximum likelihood estimation of parameters.

As we will explain in the next section, we have obtained an empirical data set of radionuclide half-lives which seems to follow a power law distribution, at least, in a partial range of the whole domain. Given that, we will try to discern, with an accurate statistical study, if the hypothesis of the power law is a plausible one for this data set. Since it is the most extended test for non normal distributions, we will use the KS statistic to test the hypothesis. In our particular study, we estimate the parameters from the same data set in which we will test the null hypothesis, ergo the statistical significance of the results can not be computed by the standard way. Therefore, the goodness-of-fit test is computed using the same method, based on Monte Carlo simulations, as proposed by Clauset *et al.* in [1].

This work is organized as follows. In chapter 3, we explain what represents in nuclear physics the half-life of a given radionuclide, how the radionuclide data has been obtained and some interesting images that illustrate and motivate our research work. In the next chapter, we introduce the power law distribution, distinguishing the cases of the pure power law and the truncated power law, and the truncated log-normal distribution. In chapter 5, we describe the probability and statistical methodologies used for the distribution parameters estimation and the goodness-of-fit performance, and the main algorithm of the code implementing it. The discussion of the results is found in chapter 6, which is divided in three parts where we show the results for each distribution analyzed in the study. Finally, chapter 7 is intended to summarize and discuss the results in a general manner and present the main conclusions we can extract from this work.

3 The Nuclide Data

In nuclear physics the **half-life** for a given radionuclide is the time for half of radionuclides in any sample to undergo radioactive decay. After two half-lives, there will be one fourth the original sample, after three half-lives one eight the original sample, and so forth. In other words, the radioactive half-life gives a pattern of reduction to half in any successive half-life period. So, the mathematical formula for the pattern is

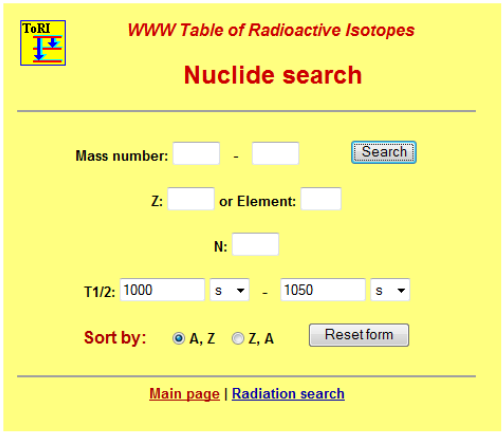
$$N = N_0 2^{-t/T_{1/2}} = N_0 e^{-t \ln 2/T_{1/2}} = N_0 e^{-\lambda t}$$

where N_0 would be the initial number of the radionuclides in the sample, N the resulting number of radionuclides after a time t, $T_{1/2}$ the characteristic half-life of the radionuclide, λ the characteristic decay constant and $1/\lambda$ the average lifetime. The formula above is the widely known radioactive disintegration law.

The rate of radioactive decay is typically expressed in terms of either the radioactive half-life, or the radioactive decay constant. Both give the same information, so either may be used to characterize the decay of a given radionuclide.

Since we are interested in the distribution of the half-lives of the radionuclides we need information from all possible isotopes known. The Lund/LBNL Nuclear Data Search is a web page from Lawrence Berkeley National Laboratory and the Department of Physics of the Lund University from Sweden, where one can find full information about nuclear structure and decay data (http://nucleardata.nuclear.lu.se/nucleardata/toi). For our interests, we have used a data table with information about the half-life, mass number, atomic number and other relevant physical magnitudes of each known isotope. The figure 2.a shows the appearance of the web data table and figure 2.b an example of searching radionuclides with half-lives from 1000 to 1050 seconds.

a)



b)

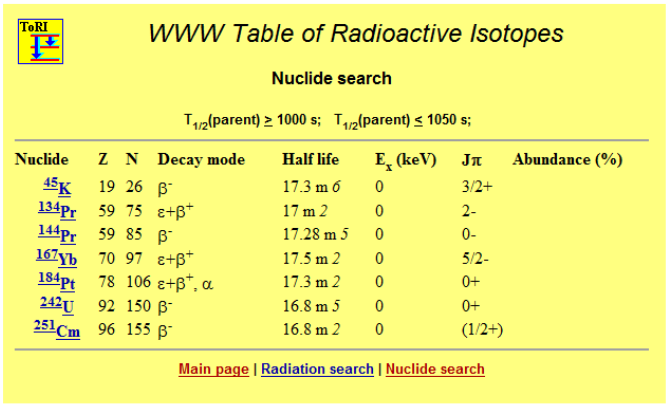


Figure 2: (a) Table where we have obtained the half-lives of the readionuclides. Since we were interested in obtaining the maximum number of radionuclides we put in the field $T_{1/2}$ the values from $1 \cdot 10^{-34}$ to $1 \cdot 10^{34}$ seconds, and no more restrictions to the search. The nuclides with half-life up to $1 \cdot 10^{34}$ s are considered stable. (b) This is an example of the search that one can perform in the web data table to find radioactive isotopes with half-life value between 1000 and 1050 seconds, both included.

By means of this table, we have obtained a data set consisting in 3001 unstable nuclei ranging from ^{3}H (Z=1; N=2) to ^{269}Hs (Z=108; N=161). The number at the left of the nuclide noun represents the mass number of the nuclide, which is the sum of the number of neutrons (N) plus the number of protons (Z). In figure 3 we show a map of N Vs. Z for our data set, where each point in the map represents a given radionuclide of the set. The blank straight line in the map is where one could find all the stable nuclides.

In addition, we have included the unstable nuclide ^{209}Bi (see reference [6] for further information) which can not be found in the table since their properties

were studied after the last actualization of the table. In sum, we have collected a list of 3002 unstable nuclides.

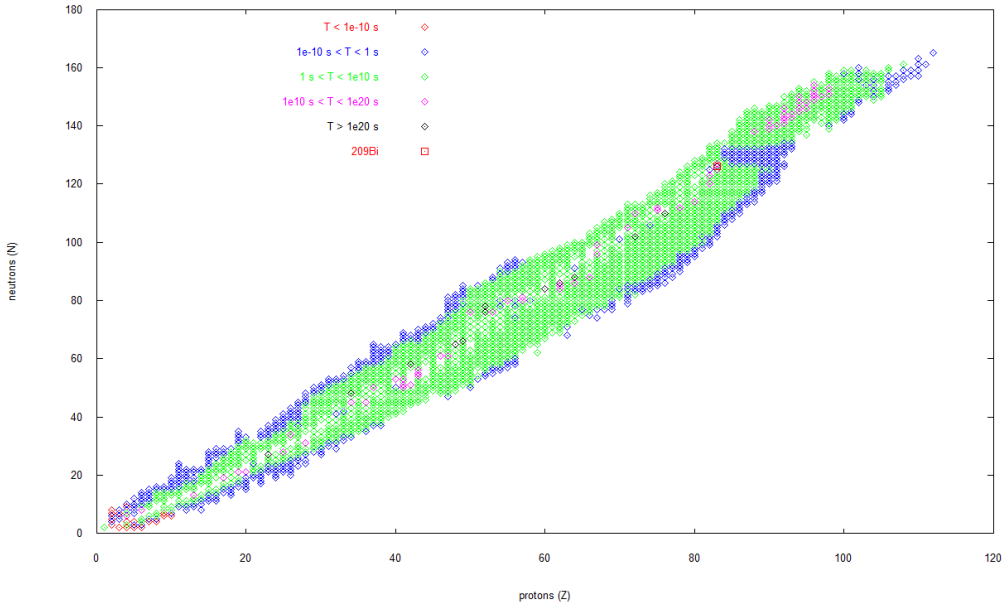


Figure 3: Representation of the radioisotopes data set on an N Vs. Z map. The red square shows the position of the ^{209}Bi isotope. The map is plotted in different colours to separate the radionuclides in five different groups depending on the value of their half-life (see the legend).

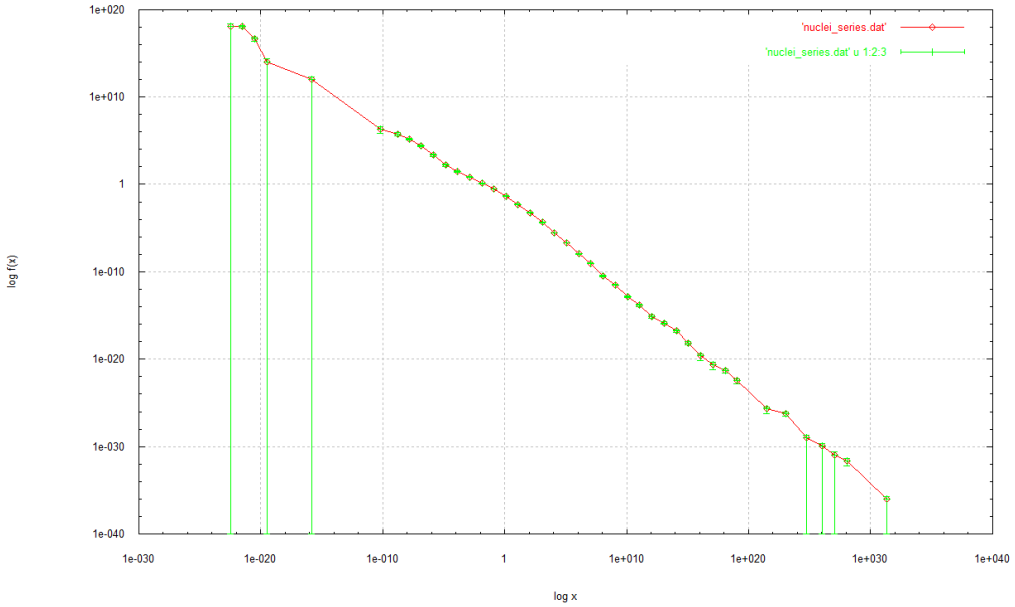


Figure 4: This is the empirical probability density of the data after a double logarithmic rescaling, calculated with the method of logarithmic binning—see [5]. The large error bars at the end of the plot and in the beginning indicates the presence of only one data point in the specific bin, so the error it is, in fact, not defined (see [7] for the evaluation of the error).

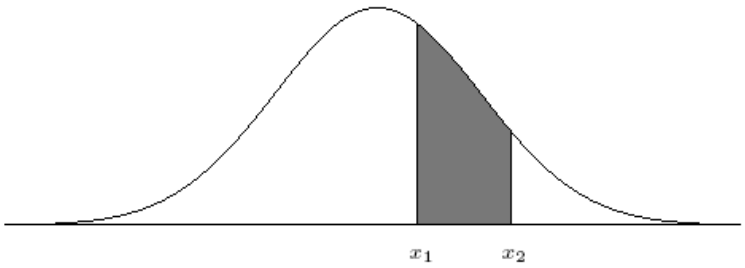
One immediately realizes that the half-lives vary across many orders of magnitude, being the shortest one $3.04 \cdot 10^{-22}$ seconds for the ^{5}Li and the largest one $6.94 \cdot 10^{31}$ seconds for the ^{128}Te. This large range of variation provides an interesting problem for the calculation of statistical properties. Figure 4 illustrates the empirical density distribution of the complete data set.

4 Continuous Distributions

The continuous probability distribution of a given continuous random variable X is completely determined by its probability density, commonly denoted by $f(x)$, which is a non-negative function with $\int_{\mathbb{R}} f(x)dx = 1$. The basic idea is that probabilities are defined by areas under the graph of $f(x)$. That is, a random variable X has density $f(x)$ if, and only if, for all $x_1 \leq x_2$, $x_1, x_2 \in \overline{\mathbb{R}}$,

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx,$$

which is the area shaded in the following diagram:



A continuous random variable with a power-law distribution has a probability density of the form

$$\boxed{\begin{array}{ll} f(x) = Cx^{-\alpha}, & C > 0 \\ \forall x \in [a, b] & \alpha > 0,\ a > 0,\ a < b \in \overline{\mathbb{R}} \end{array}} \tag{1}$$

being α a positive constant parameter known as the *exponent* or *scaling parameter* of the distribution and C the normalization constant, which depends on the value of α. The parameters a and b are an upper and a lower bound for the distributions. In this work we will separate the power law distributions in two different subsets (since the mathematical and computational treatments are not exactly the same): power laws in a semi-infinite interval $[a, \infty)$ and power laws in a finite interval $[a, b]$.

In the subsections below we will perform a little introduction to each distribution mentioned above and a brief introduction to the truncated log-normal distribution, given that we suspect the data follows this distribution in some regions of the whole domain.

4.1 Power-law when $x \in [a, \infty)$

In the case of considering a power-law distribution in the interval $[a, \infty)$, one can easy find the constant C in formula (1) by the normalization requirement

$$1 = \int_a^\infty f(x)dx = C \int_a^\infty x^{-\alpha}dx = \frac{C}{1-\alpha}\left[x^{-\alpha+1}\right]_a^\infty . \tag{2}$$

One can observe that the equation is only possible for $\alpha > 1$, since otherwise the integral in (2) diverges and the requirement of integral in the whole domain of the density equal to 1 is not satisfied. Therefore, if $\alpha > 1$ then expression (2) gives

$$C = (\alpha - 1)a^{\alpha-1}, \tag{3}$$

and the expression for the density itself is

$$f(x) = \frac{\alpha - 1}{a^{1-\alpha}}\frac{1}{x^\alpha}. \tag{4}$$

Some distributions follow a power law for part of their range but are cut off at high values of x. In some cases, above some value they deviate from the power law and fall off quickly towards zero. If this happens, then the distribution may be normalizable no matter what the value of the exponent α. Anyway, in the next subsection we introduce the truncated power law distributions without wondering what is the behaviour after or before the power law range.

4.2 Power-law when $x \in [a, b]$, $b \in \mathbb{R}$

It is convenient for our work to define a power law distribution in the case of having both a lower and an upper bound for the values of x. Labeling the lower and the upper bound as a and b, respectively, one can normalize the density (1) in the same way as we did in the previous case. Then, the normalization constant is, $\forall \alpha > 0$,

$$C = \frac{1-\alpha}{b^{1-\alpha} - a^{1-\alpha}} \tag{5}$$

and we obtain the exact formula for the density of a power law in the finite interval $[a, b]$:

$$f(x) = \frac{\alpha - 1}{a^{1-\alpha} - b^{1-\alpha}}\frac{1}{x^\alpha}. \tag{6}$$

We can see that doing $b \to \infty$ in this formula we find again the probability density (4) for the case $[a, \infty)$ if $\alpha > 1$.

The next step is to introduce the third candidate of distribution we think that can be useful in the modelization of the behaviour of the radionuclide half-lives data.

4.3 Normal distribution when $x \in [A, B]$

We say that X is a normal random variable (or simply that X is normally distributed) if the probability density of X is given by

$$\boxed{\begin{array}{ll} f(x) = Ce^{-\frac{(x-\mu)^2}{2\sigma^2}}, & C > 0 \\ \forall x \in \mathbb{R} & \mu \in \mathbb{R},\ \sigma > 0 \end{array}} \tag{7}$$

where μ is the mean of x, σ the standard deviation and C the normalization constant. Typically, the value for the constant C is $\frac{1}{\sqrt{2\pi\sigma^2}}$ given that the domain of the density is $(-\infty, \infty)$. However, we are interested in the cases where the domain of the distribution is truncated. Then, if A and B are the lower and the upper bound of the truncated domain, the interval in which we have to work becomes $[A, B]$, where $A \in \mathbb{R}$ and $A < B \in \mathbb{R}$. As we did in the case of truncated power law, we have to compute the correspondig C taking into account that the density out of the interval is zero. After some algebra, the constant C takes the value

$$C = \frac{2}{\sqrt{2\pi\sigma^2}} \left[\operatorname{erf}\left(\frac{B-\mu}{\sqrt{2\sigma^2}}\right) - \operatorname{erf}\left(\frac{A-\mu}{\sqrt{2\sigma^2}}\right) \right]^{-1} \tag{8}$$

and the exact form for the probability density is

$$f(x) = \frac{2}{\sqrt{2\pi\sigma^2}} \left[\operatorname{erf}\left(\frac{B-\mu}{\sqrt{2\sigma^2}}\right) - \operatorname{erf}\left(\frac{A-\mu}{\sqrt{2\sigma^2}}\right) \right]^{-1} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{9}$$

where $\operatorname{erf}(\cdot)$ is the known error function given by

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

But, as said in the beginning of the chapter, we think the data following not a normal distribution but a log-normal distribution. Even though, a variable with a log-normal distribution is a variable whose logarithm is normally distributed. Then, if y is a variable with a log-normal distribution, defining the simple regular change of variable $x = \ln y$ the resulting variable x becomes

normally distributed. Therefore, since it is easier to work with normal distribution than a log-normal distribution, we can take the logarithm of the empirical data and, then, we may work with it as a normally distributed sample.

Some chapters below, we will need the formula of the log-normal density. Then, by the equality

$$g(y) = f(x) \left| \frac{dx}{dy} \right|,$$

where $g(y)$ represents the log-normal density, we obtain that

$$g(y) = C \, \frac{1}{y} \, e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} \tag{10}$$

where $1/y$ comes from the factor $|dx/dy|$ and C is the constant (8). It is necessary to clarify that the domain of the variable y (our original data) is $[a, b]$ where $0 < a < b$ and $a, b \in \mathbb{R}$. Then the correspondence between the domain of the normal distribution and the domain of the log-normal distribution comes from $A = \ln a$ and $B = \ln b$.

5 Data Analysis

This chapter is pretended to describe in detail the statistical and computational issues we have required in our study, and it is divided in two parts. In the first part one can find the explanation of the three main blocks of the statistical concepts used in our work and the standard method of simulating data from a known probability distribution. In the second part there is summary of the main algorithm implemented in the codes for doing the data analysis.

5.1 Tools for the data analysis

Given the nature of a power law distribution, a commonly method used in the estimation of the scaling parameter of the distribution is performing a least-squares linear regression, or other variations on the same theme. This implies the binning of the data in small intervals and count the number of observations in each bin, generating a histogram that reproduces empirically the probability density of the distribution. After a doubly logarithmic rescaling one obtains a straight line whose slope is pretended to be the exponent of the power law.

Unfortunately, this method generates substantial inaccurate results since it assumes independent Gaussian distributed errors in the bins of the histogram but it is not applicable after the logarithmic rescaling of the plot—see [1]. Recent studies on estimating the exponent of power law distributions support this fact and confirm that maximum likelihood estimation outperforms other methods in both accuracy and precision—see [4].

5.1.1 Maximum Likelihood parameter Estimator (MLE)

Definition 5.1. *If $x_1, \ldots, x_N$ are the values of random sample of a variable X with density $f_X(x;\theta)$, where θ is the characteristic parameter of the density, the likelihood function of this random sample, is given by*

$$L(\theta) = f_{X_1,\ldots,X_N}(x_1, \ldots, x_N; \theta) \tag{11}$$

for the values of θ in a given domain. In this case $f_{X_1,\ldots,X_N}(x_1, \ldots, x_N; \theta)$ is the value of the joint probability density of the random variables $X_1, \ldots, X_N$ in $x_1, \ldots, x_N$.

The maximum likelihood method consist in maximizing the likelihood function respect to θ—see [8]. More explicitly, given the assumption of independent

and identically distributed data in the random sample, the function $L(\theta)$ to maximize becomes

$$L(\theta) = \prod_{i=1}^{N} f_{X_i}(x_i;\theta) \tag{12}$$

where x_i represents our data points and $f_{X_i}(x_i;\theta)$ are, obviously, identical versions of $f_X(x;\theta)$. Then, the equation we have to solve to find the maximum of $L(\theta)$ is the following

$$\frac{\partial L(\theta)}{\partial \theta} = 0. \tag{13}$$

We will refer to the value $\hat{\theta}$ that solves this equation as the maximum likelihood estimator (MLE) of θ.

It will be convenient for our job to work with log-likelihood function $\mathcal{L}(\theta)$ instead of the likelihood function since the logarithm properties will make easy the finding of MLE $\hat{\alpha}$ in the case of power laws and $\hat{\sigma}$ and $\hat{\mu}$ in the log-normal distribution case. The log-likelihood function is defined by taking the logarithm of the likelihood function and dividing by the total number of data points N, then

$$\mathcal{L}(\theta) = \frac{1}{N} \ln L(\theta), \tag{14}$$

and the equation we will solve to find the maximum it is the same as before but changing $L(\theta)$ for $\mathcal{L}(\theta)$. It is easy to see that the maximum of the likelihood function and the maximum of log-likelihood function are the same.

In the case of a power law in the domain $[a, +\infty)$, the log-likelihood function to maximize is

$$\mathcal{L}(\alpha) = \ln(\alpha - 1) - \ln a - \frac{\alpha}{N} \sum_{i=1}^{N} \ln\left(\frac{x_i}{a}\right). \tag{15}$$

Setting $\partial\mathcal{L}(\alpha)/\partial\alpha = 0$ and after some algebra one can find that the exact form for the MLE of the scaling parameter is

$$\hat{\alpha} = 1 + N \left[\sum_{i=1}^{N} \ln\left(\frac{x_i}{a}\right)\right]^{-1}. \tag{16}$$

In the case of a power law in the domain $[a, b]$, the log-likelihood function takes the form

$$\mathcal{L}(\alpha) = \ln\left(\frac{1-\alpha}{b^{1-\alpha} - a^{1-\alpha}}\right) - \frac{\alpha}{N} \sum_{i=1}^{N} \ln x_i. \tag{17}$$

Unfortunately, the derivative of (17), as something usual in this context, it doesn't have analytical solution when set equal to zero, so we have to calculate the maximum value $\hat{\alpha}$ numerically [1].

Finally, in the case of data following a log-normal distribution in $[a, b]$, we have two parameters to estimate, σ and μ. Using expression (9) it is clear that the likelihood function in this case will be

$$\mathcal{L}(\mu, \sigma) = \ln C(\mu, \sigma) - \frac{1}{2\sigma^2 N} \sum_{i=1}^{N} (x_i - \mu)^2 \tag{18}$$

where $C(\mu, \sigma)$ comes from the normalization constant (8). Given the complexity of the function we will not compute the likelihood equations and, as in the previous case, we will use a numerical recipe to find the MLE for μ and σ, directly working with the log-likelihood function.

5.1.2 Kolmogorov-Smirnov statistic

For our purposes it will be necessary to estimate how far are the theoretical models we find from the empirical data. There are a variety of measures for quantifying the distance between two probability distributions, but for non-normal data the commonest is the Kolmogorov-Smirnov test or KS statistic based on the Glivenko-Cantelli theorem (a fundamental theorem of statistics—see [9]), which is simply the maximum distance between the cumulative distribution function (or their complementary—see section 9.1 of the Appendix for further information about the complementary CDF functions used) of the data and the theoretical model. The KS statistic is given by

$$d_{KS} = \sup_{a \leq x \leq b} |S_N(x) - F(x; \theta)| \tag{19}$$

where $S_N(x)$ is the empirical CDF of the experimental data (see section 9.2 of the Appendix for information about computing it) and $F(x; \theta)$ is the CDF of the theoretical model in the region $a \leq x \leq b$. In our case, the theoretical model is the model with the parameters obtained using the maximum likelihood method (i.e. the model with $\hat{\theta}$). Then, we redefine the KS statistic as

$$\hat{d}_{KS} = \sup_{a \leq x \leq b} \left| S_N(x) - \hat{F}(x; \hat{\theta}(x)) \right|. \tag{20}$$

[1] We use here a subroutine (`amoeba.for`) from [10] that implements the downhill simplex method for minimizing functions.

A property of the KS statistic that makes it very useful is that its distribution in the case of data sets drawn from the theoretical distribution (the null hypothesis) can be calculated, at least to useful approximation, thus giving the significance of any observed nonzero value of d_{KS}. Unfortunately, we can not use this interesting property since the KS statistic is not stable under theoretical models obtained by estimation of parameters. Instead of that, we shall do Monte Carlo simulations to calculate the significance of our results. This will be the aim of the next subsection.

5.1.3 Simulated construction of $\hat{d}_{KS}$ distribution

As previously explained, we can not use directly the distribution of the KS statistic to give a significance of our results because we use the maximum likelihood method to fit the best theoretical model for our data. Given that, to estimate the statistical significance of the calculated KS statistics we compute a p-value that is evaluated following a Monte Carlo procedure.

This procedure consists in simulating n synthetic data (say 1000 simulations), each of these with N random numbers following the theoretical model we found by maximum likelihood estimation of parameters, given by $\hat{\theta}$. After that, by maximum likelihood again, we calculate the theoretical model (if we do 1000 simulations, we do it 1000 times) and we compute the KS statistic of each simulation. Mathematically,

$$\hat{d}^{\text{sim}}_{KS,i} = \sup_{a \leq x \leq b} \left| S^{\text{sim}}_i(x) - \hat{F}^{\text{sim}}_i(x; \hat{\theta}(x)) \right| \qquad i = 1, \ldots, n \tag{21}$$

where $S^{\text{sim}}_i(x)$ is the CDF of the simulated data i and $\hat{F}^{\text{sim}}_i(x; \hat{\theta}(x))$ is the theoretical model CDF relative to each simulated data i. Finally we compute the p-value simply counting the $\hat{d}^{\text{sim}}_{KS,i}$ cases bigger than the $\hat{d}_{KS}$ value for the empirical data, against total cases. In other words,

$$p = \frac{\text{Number of simulations}(\hat{d}^{\text{sim}}_{KS,i} > \hat{d}_{KS})}{n}. \tag{22}$$

We can define the simulated p-value as the probability of obtaining a $\hat{d}_{KS}$ larger than the one that was actually observed, assuming that the null hypothesis is true. In general, it is straightforward to see that, under the null hypothesis, with the domain $[a, b]$ fixed, the p-value itself will be randomly distributed, with a uniform distribution in $[0, 1]$.

The p-value gives us a measure of how far is our adjustment from the simulations randomly generated. For example, if $p = 0.6$ means that the 60%

of KS distances calculated from random simulations are bigger than $\hat{d}_{KS}$ from empirical data, then we can not reject the null hypothesis because, in some way, this distance $\hat{d}_{KS}$ could be generated by random simulation from the theoretical model. Conversely, obtaining a very small p-value (say $p = 0.0001$) means that our KS distance is bigger than the majority of the KS distances from all the simulations and then the error introduced in the adjustment of the data seems to be bigger than the random error of the simulations. Then, we must reject the null hypothesis.

An important question is how to decide how many simulations we need to generate. Since we want approximately an accuracy of $\epsilon = 0.01$ in our estimates of the p-value we will do over $\frac{1}{4}\epsilon^{-2}$ simulations each time, according to [1].

5.1.4 Simulations

The most commonly used method for generating random numbers of non-uniform distributions is the *transformation method*—see Appendix 9.3 and reference [11]. The method involves in fact that the CDF of a given random variable follows a uniform $[0, 1]$ distribution by itself. Then, if we know the cumulative distribution function for a given variable X then

$$F_X(x) \sim \text{unif}[0, 1] \tag{23}$$

and inverting this formula we obtain

$$X \sim F_X^{-1}\left(\text{unif}[0, 1]\right). \tag{24}$$

In conclusion, if we can compute the inverse of the distribution function of a given random variable X we could generate a sequence of random numbers following the distribution of X by applying the inverse of the CDF to a sequence of uniform random numbers. In the case of a normal distribution, the CDF contains the error function, and its inversion is not easily available in practice. However, a generalization of the method to two dimensions allows the use in this case of cylindrical coordinates, from which the rectangular components can be immediately obtained. This is the Box-Mueller-Wiener algorithm—see [10].

Distribution	Random Numbers
Power law in $[a, \infty)$	$x = \frac{a}{(1-u)^{\frac{1}{\hat{\alpha}-1}}}$
Power law in $[a, b]$	$x = \frac{a}{\left\{1-\left[1-\left(\frac{a}{b}\right)^{\hat{\alpha}-1}\right]u\right\}^{\frac{1}{\hat{\alpha}-1}}}$
Log-Normal	$x = e^{(\hat{\mu}+\hat{\sigma}g)}$

Table 1: In the first two cases u represents a uniform random number obtained using the `ran3.for`, a uniform random number generator from [10]. In the third case the number g is a Gaussian random number with variance equal to 1 and 0 mean. This number g is obtained by means of Box-Mueller-Wiener algorithm, as explained in the main text. Parameters $\hat{\mu}$ and $\hat{\sigma}$ are the characteristic parameters of the log-normal distribution, obtained by maximum likelihood method.

5.2 Procedure for the data analysis

There exist multiple commercial software for general data analysis and statistics. But, in our case, given the very specific skills of the topics we have treated, it has been necessary to implement all the routines for the data analysis by ourselves. The main goal of this subsection is to explain the procedure we have followed, in the F77 routines we have created, to analyze the nuclide data. It is important to remark that we have used several subroutines and functions from [10] in our F77 codes.

For simplicity, we have implemented separately the codes for the three cases of distributions we have studied. The main procedure is the same in the three cases and their scheme is the following:

1. **Read** and **sort** the nuclide data in increasing order.

2. **Loop**: Take a lower bound a_i for the data, $i = 1, \ldots, N_a$.

   (a) **Loop**: Take an upper bound b_j for the data, $j = 1, \ldots, N_b$ [2].

      i. Compute the value $\hat{\alpha}_{ij}$, or $\hat{\mu}_{ij}$ and $\hat{\sigma}_{ij}$ in the case of log-normal distribution, by **MLE** method.

      ii. Compute the **KS statistic** ($\hat{d}_{KS}$) for the interval.

      iii. Compute the **p-value**:

         A. Simulate a distribution with parameters: a_i, b_j, $\hat{\alpha}_{ij}$ or $\hat{\mu}_{ij}$ and $\hat{\sigma}_{ij}$, and n_{ij} (the number of data points in $[a_i, b_j]$).

         B. Compute a new $\hat{\alpha}_{\text{sim}}$ (or $\hat{\mu}_{\text{sim}}$ and $\hat{\sigma}_{\text{sim}}$) with MLE, now with the simulated data points.

         C. Calculate the KS statistic for the simulation:

$$\hat{d}^{\text{sim}}_{KS} = \sup_{a \le x \le b} \left| S^{\text{sim}}(x) - \hat{F}^{\text{sim}}(x; \hat{\theta}(x)) \right|,$$

            where $S^c_{\text{sim}}(x)$ and $F^{\text{sim}}(x; \hat{\theta}(x))$ are the CDF of the simulated data and the CDF of the theoretical model that best fits the simulated data (with $\hat{\theta}(x) = \hat{\alpha}_{\text{sim}}$, or $\hat{\theta}(x) = (\hat{\mu}_{\text{sim}}, \hat{\sigma}_{\text{sim}})$, from B), respectively.

         D. Check if $\hat{d}^{\text{sim}}_{KS} > \hat{d}_{KS}$. Go to A.

[2] In the case of power law in $[a, \infty)$ we omit this step

E. Repeat the three steps before 2500 times for each a_i, b_j, $\hat{\alpha}_{ij}$ (or $\hat{\mu}_{ij}$ and $\hat{\sigma}_{ij}$) and n_{ij}, and calculate:

$$p = \frac{\text{Number of } (\hat{d}_{KS}^{\text{sim}} > \hat{d}_{KS})}{2500}.$$

(b) **Go to** (a). **End** when $b = b_{N_b}$.

3. **Go to** 2. **End** when $a = a_{N_a}$.

4. **End** program.

Other F77 routines have been used in order to pre-processing the nuclide data from the website, to extract information from data, etc. But we will not explain here the details since this is not the main object of the work.

6 Results

This section is organized in four parts. The three first are destined to show and discuss separately, for the three approaches treated in this work, the results obtained after the application of the procedures described to our data. Finally, in the last one, we summarize all the results and we make a global discussion of them.

6.1 Power-law $x \in [a, \infty)$

As explained in subsection 5.2 we have done 2500 simulations for each fixed value of the lower bound a. We have moved a from $1.47 \cdot 10^{-5}$ to $6.81 \cdot 10^{21}$ seconds with the formula $a_i = a_0 10^{1/6} 10^{2i/3}$ for $i = 0, \ldots, 40$ and $a_0 = 10^{-5}$. The reason of the factor $10^{1/6}$ is in order to avoid integer powers of 10, to which the data have tendency to be rounded. We show the results in table 2. One can see in this table that we have omitted the results of the simulations with lower bounds less than $1.47 \cdot 10^{-1}$ and greater than $3.16 \cdot 10^{14}$. In the first case the reason is simply that the p-values for these simulations are zero and they do not contribute in any relevant information for our study. In the second case we have considered that the number N of data points is too small to be statistically significant and they do not contribute in any important information too.

With the results in table 2 we can extract interesting conclusions, not only for the nuclide data, but for the Clauset method too. It is clear that does not exist a pure power law, at least for a large range of the data. However, we show in table 2 that, if there exist, the power law must have a lower bound around $a = 1.47 \cdot 10^9$ seconds since the simulations for this range provide an acceptable p-value ($p = 60\%$), despite that the number of data N seems to be a little bit poor ($N = 97$). Values of a up to $1.47 \cdot 10^7$ s also provide reasonable p-values.

It is important to note that the best p-value in the table do not correspond with the lowest KS distance. As one can observe, the lowest value of the KS distance ($d_{KS} = 0.038$) is obtained when the lower bound takes the value $1.47 \cdot 10^1$ seconds and the p-value in this case is 0. This fact reveals that the recipe proposed for Clauset et al. in [1] to analyze power law distributed data has to be carefully used.

In figure 5 we show the fits discussed above. In the case of the fit that gives the greatest p-value the exponent of the power law is $\alpha = 1.08$ and the case with minimal KS distance an exponent of $\alpha = 1.15$. We can note in the

a(s)	$\hat{\alpha} \pm \sigma$	$\hat{d}_{KS}$	N	p
$1.47 \cdot 10^{-1}$	1.12± 0.00	0.110	2565	0.00
$6.81 \cdot 10^{-1}$	1.13± 0.00	0.090	2317	0.00
$3.16 \cdot 10^{0}$	1.14± 0.00	0.068	2050	0.00
$1.47 \cdot 10^{1}$	1.15± 0.00	0.038	1754	0.00
$6.81 \cdot 10^{1}$	1.16± 0.00	0.040	1440	0.00
$3.16 \cdot 10^{2}$	1.16± 0.00	0.051	1119	0.00
$1.47 \cdot 10^{3}$	1.15± 0.01	0.065	853	0.00
$6.81 \cdot 10^{3}$	1.14± 0.01	0.080	644	0.00
$3.16 \cdot 10^{4}$	1.14± 0.01	0.103	502	0.00
$1.47 \cdot 10^{5}$	1.13± 0.01	0.124	372	0.00
$6.81 \cdot 10^{5}$	1.12± 0.01	0.137	287	0.00
$3.16 \cdot 10^{6}$	1.11± 0.01	0.137	223	0.00
$1.47 \cdot 10^{7}$	1.09± 0.01	0.069	161	0.19
$6.81 \cdot 10^{7}$	1.09± 0.01	0.057	133	0.57
$3.16 \cdot 10^{8}$	1.09± 0.01	0.065	116	0.47
$1.47 \cdot 10^{9}$	1.08± 0.01	0.065	97	0.60
$6.81 \cdot 10^{9}$	1.08± 0.01	0.083	84	0.36
$3.16 \cdot 10^{10}$	1.08± 0.01	0.086	76	0.38
$1.47 \cdot 10^{11}$	1.09± 0.01	0.099	71	0.25
$6.81 \cdot 10^{11}$	1.09± 0.01	0.115	62	0.16
$3.16 \cdot 10^{12}$	1.09± 0.01	0.138	56	0.07
$1.47 \cdot 10^{13}$	1.09± 0.01	0.135	45	0.15
$6.81 \cdot 10^{13}$	1.08± 0.01	0.135	38	0.23
$3.16 \cdot 10^{14}$	1.08± 0.01	0.138	30	0.36

Table 2: Estimates of the scaling parameter $\hat{\alpha}$, the KS distance ($\hat{d}_{KS}$) and the p-value (p) by increasing values of the lower bound a. The column N is the number of data points. The error on $\hat{\alpha}$ is given by the standard error σ and the error in p is 0.01.

figure how the fit with largest p-value works well at the end of the empirical density (since the lower bound is $a = 1.47 \cdot 10^9$) and the fit with minimal KS distance ($a = 1.47 \cdot 10^1$) performs better at the middle of the empirical density but worse at the end.

It has been interesting to plot the evolution of the KS distance versus the lower bound of the distribution. In figure 6 one can observe this evolution and note how the best p-value is obtained when the distance falls in a local minimum with approximate value around 0.060.

For the reasons explained above it has been convenient to fit other distributions to our data: the truncated power law.

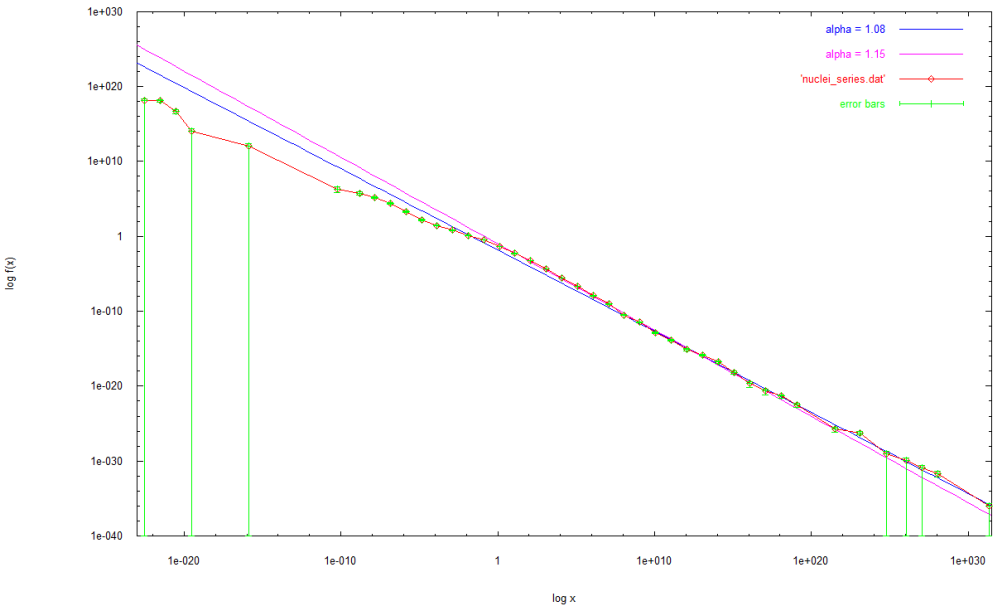


Figure 5: Here we can observe the two fits mentioned in the main text for the power law in $[a, \infty)$. In blue: the fit with $\alpha = 1.08$ that provides the maximum p-value. In purple: the fit with $\alpha = 1.15$ that provides the minimal $\hat{d}_{KS}$. The empirical distribution is the same as shown in section 3, and it is plotted in red with green error bars.

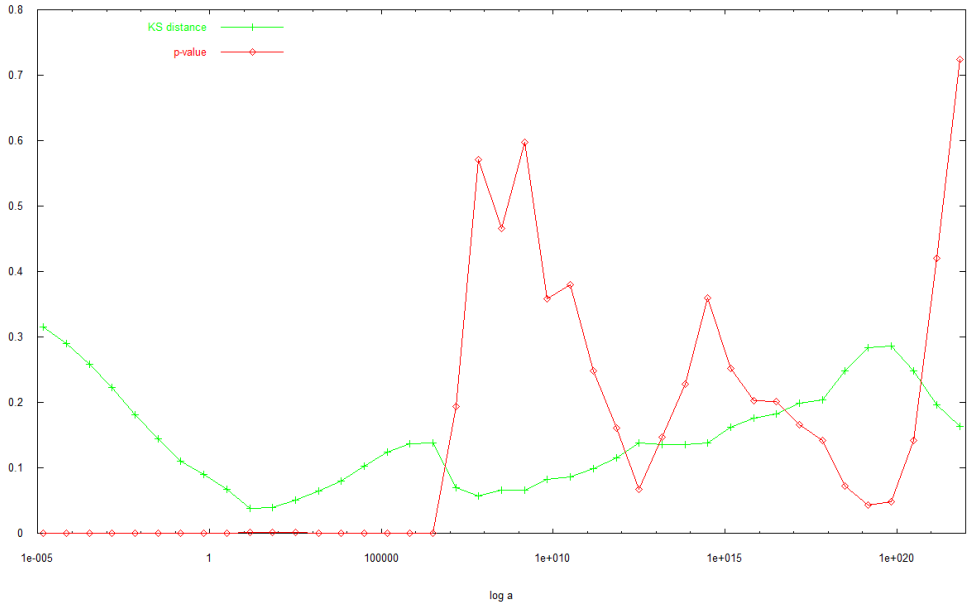


Figure 6: In green: the evolution of $\hat{d}_{KS}$ when we increase the lower bound of the case of the pure power law. One can observe the minimum ($\hat{d}_{KS}$) when $a = 1.47 \cdot 10^1$. In red: the evolution of the p-value.

6.2 Power-law $x \in [a, b]$

In order to obtain the results for the truncated power law, we have increased the values of the lower bound of the power law from $a = 1.47 \cdot 10^{-12}$ to $a = 1.47 \cdot 10^{8}$ seconds, for each value of a we do several simulations changing the upper bound b. Specifically, we have replaced the upper bound b from $1.47 \cdot 10^{10}$ to $1.47 \cdot 10^{30}$, and we have done 2500 simulations for each pair $[a_i, b_j]$. The formulas used to increment a and b have been $a_i = a_0 10^{1/6} 10^{2i/3}$ and $b_j = b_0 10^{1/6} 10^{2j/3}$ with $a_0 = 10^{-12}$ and $b_0 = 10^{10}$, respectively, and $i, j = 0, \ldots, 30$ (i.e., we have obtained $i \times j$ results since we have one p-value for each pair $[a_i, b_j]$). The most interesting results have appeared in the simulations for the value of the lower bound $a = 1.47 \cdot 10^2$. In table 3, we show the results of data analysis for this fixed value of a and some possible values of b. One can see that the largest p-value is obtained for $b = 1.47 \cdot 10^{14}$.

First of all, we can assert the existence of a power law behaviour in the range $[1.47 \cdot 10^2, 1.47 \cdot 10^{14}]$ where we have found the minimum of the KS minimum distances ($\hat{d}_{KS} = 0.016$) and the greater p-value ($p = 83\%$). The number of nuclides N for this result is 1252. In a second term, we find a reasonable values in the range $[1.47 \cdot 10^2, 3.16 \cdot 10^{21}]$ where the KS minimum distance becomes $\hat{d}_{KS} = 0.020$ and the p-value takes the not inconsiderable rate of $p = 51\%$. This last interval has the benefit of containing more nuclides than the first interval: 1271 versus 1252 (about 1.5% more).

As did in the previous subsection, we have represented in figure 8 the evolution of the KS statistic as a function of a and b, which describes a surface in the space. Then, in figure 9 we show the projection of this surface in the plane $(\ln a, \hat{d}_{KS})$.

Comparison of the exponents of the truncated power law and the pure power law reported above (1.18 versus 1.08) suggest that may be the rightmost tail of the distribution (the highest half-lives) suffers from some kind of over sampling. For some reason, it is possible that nuclei in the limit of stability have attracted more experimental interest and this makes the other nuclei to be under-represented in the sample.

In addition, since we have observed a change of slope in the empirical density, approximately when the half-life is around 1 second, we have proceeded to analyze the data before this value (range of data with low half-lives). Now, the formulas used to move the $[a_i, b_j]$ interval have been $a_i = a_0 10^{1/6} 10^{i/4}$ and $b_j = b_0 10^{1/6} 10^{j/4}$, where $i, j = 1, \ldots, 20$. The values of the upper and the lower bound where we have begun to analyze the power law have been $a_0 = 10^{-11}$ and $b_0 = 10^{-7}$.

b(s)	$\hat{\alpha} \pm \sigma$	$\hat{d}_{KS}$	N	p
$1.47 \cdot 10^{10}$	1.19 ± 0.01	0.024	1206	0.26
$6.81 \cdot 10^{10}$	1.19 ± 0.01	0.023	1211	0.34
$3.16 \cdot 10^{11}$	1.19 ± 0.01	0.024	1221	0.32
$1.47 \cdot 10^{12}$	1.18 ± 0.01	0.036	1224	0.02
$6.81 \cdot 10^{12}$	1.19 ± 0.01	0.024	1233	0.26
$3.16 \cdot 10^{13}$	1.19 ± 0.01	0.019	1240	0.60
$1.47 \cdot 10^{14}$	1.18 ± 0.01	**0.016**	1252	**0.83**
$6.81 \cdot 10^{14}$	1.18 ± 0.01	0.017	1257	0.76
$3.16 \cdot 10^{15}$	1.18 ± 0.01	0.023	1260	0.33
$1.47 \cdot 10^{16}$	1.18 ± 0.01	0.018	1261	0.68
$6.81 \cdot 10^{16}$	1.18 ± 0.01	0.022	1263	0.35
$3.16 \cdot 10^{17}$	1.18 ± 0.01	0.022	1264	0.36
$1.47 \cdot 10^{18}$	1.18 ± 0.01	0.021	1267	0.47
$6.81 \cdot 10^{18}$	1.17 ± 0.00	0.029	1270	0.10
$3.16 \cdot 10^{19}$	1.17 ± 0.00	0.028	1271	0.13
$1.47 \cdot 10^{20}$	1.17 ± 0.00	0.029	1271	0.11
$6.81 \cdot 10^{20}$	1.17 ± 0.00	0.029	1271	0.11
$3.16 \cdot 10^{21}$	1.18 ± 0.00	0.020	1271	0.51
$1.47 \cdot 10^{22}$	1.17 ± 0.00	0.028	1273	0.15

Table 3: From right to left: the upper bound of the half-lives in seconds, the estimation of the scaling parameter $\hat{\alpha}$ (with associated statistical error σ—see reference [12]), the value of the KS statistic, the number N of data points between a and b and the p-value for the simulations with $a = 1.47 \cdot 10^2$.

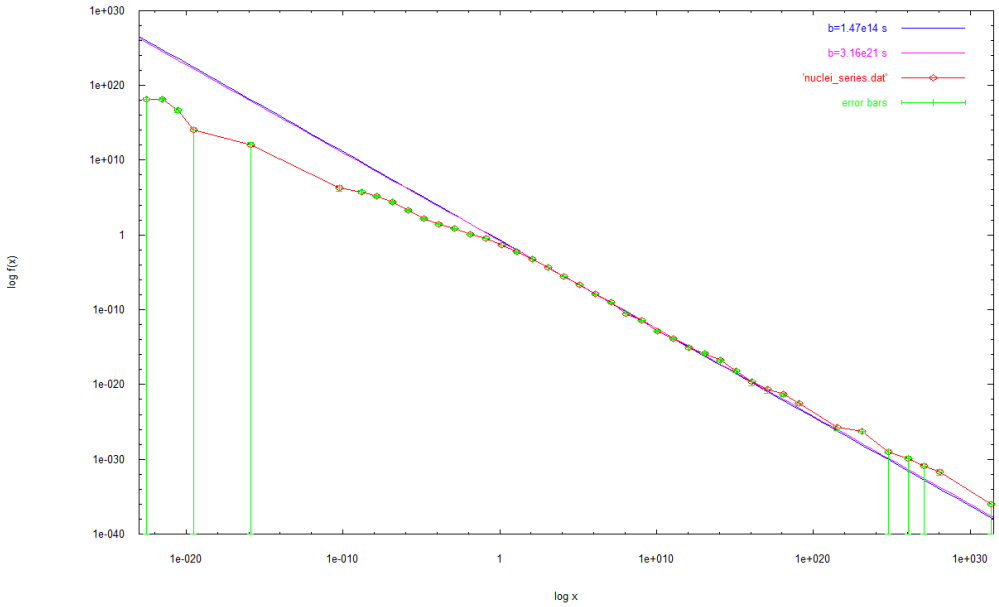


Figure 7: In blue: the fit for the range $[1.47 \cdot 10^2, 1.47 \cdot 10^{14}]$. In purple: the fit for the case of range $[1.47 \cdot 10^2, 3.16 \cdot 10^{21}]$.

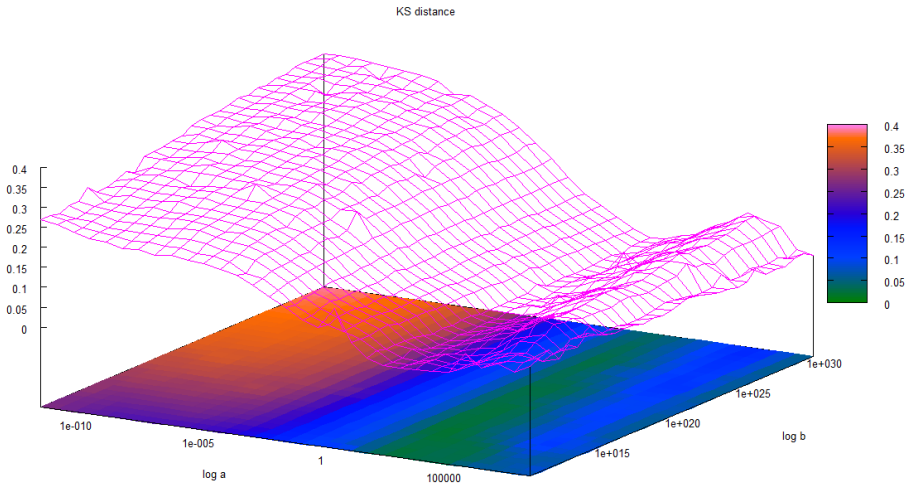


Figure 8: Surface described by the KS statistic, as a function of $\ln a$ and $\ln b$.

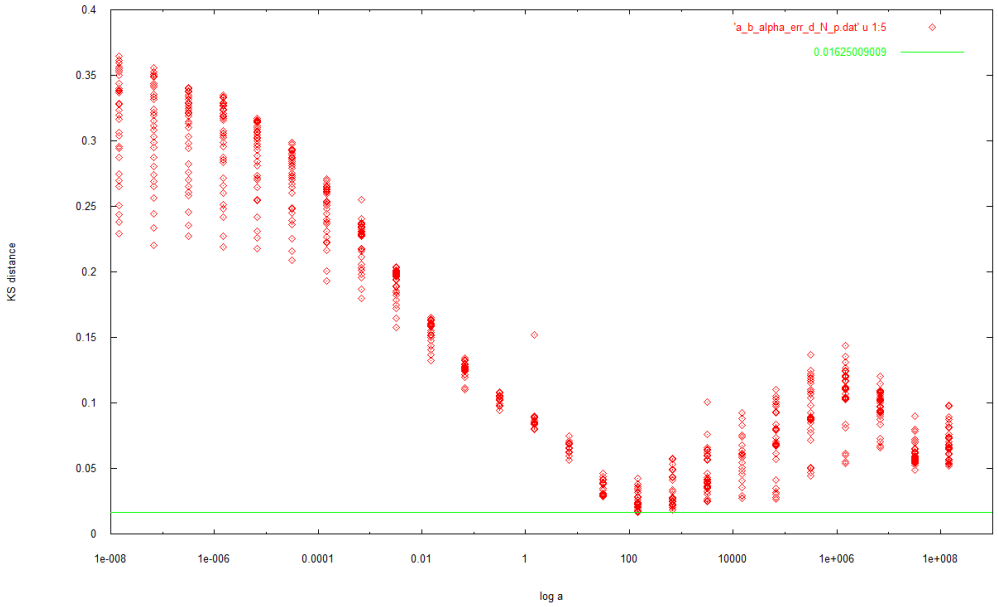


Figure 9: Projection of the surface described by $\hat{d}_{KS}(\ln a, \ln b)$ in the plane $(\ln a, \hat{d}_{KS})$.

b(s)	$\hat{\alpha} \pm \sigma$	$\hat{d}_{KS}$	N	p
$1.47 \cdot 10^{-7}$	0.82±0.16	0.062	29	1.00
$2.61 \cdot 10^{-7}$	0.86±0.13	0.060	34	0.99
$4.64 \cdot 10^{-7}$	0.93±0.11	0.099	37	0.52
$8.25 \cdot 10^{-7}$	0.86±0.09	0.078	46	0.73
$1.47 \cdot 10^{-6}$	0.88±0.08	**0.055**	51	**0.97**
$2.61 \cdot 10^{-6}$	0.95±0.07	0.093	53	0.35
$4.64 \cdot 10^{-6}$	0.97±0.06	0.093	57	0.27
$8.25 \cdot 10^{-6}$	0.99±0.06	0.099	60	0.18
$1.47 \cdot 10^{-5}$	1.02±0.05	0.099	62	0.15
$2.61 \cdot 10^{-5}$	1.00±0.05	0.088	68	0.24
$4.64 \cdot 10^{-5}$	0.99±0.04	0.087	73	0.21

Table 4: This table shows the results for the adjustment of a truncated power law with lower bound $a = 2.61 \cdot 10^{-9}$ seconds and upper bound form $1.47 \cdot 10^{-7}$ to $4.64 \cdot 10^{-5}$ seconds. The smallest KS distance and corresponding p-value are in bold. We show the same parameters as in table 3.

One can realize in table 4 that all the p-values are greater than 0.10. Since we have fixed the barrier of p in this value, to rule out the null hypothesis, we must accept the null hypothesis in all cases shown in this table. Even though, according to [1], the interval $[2.61 \cdot 10^{-9}, 1.47 \cdot 10^{-6}]$ with $\hat{\alpha} = 0.88$ has the property of minimizing the KS distance and, therefore, we will take it as the best fit of our data in the range of low half-lives. However, it is important to remark that the poor number of data points at low half-lives can affect negatively at the results shown in table 4. Perhaps, further research about new radionuclides with low half-life could provide more information in the future to improve these hypothesis.

It is interesting to note that, for high values of the half-lives, there is a power law behaviour with exponent $\hat{\alpha}_H = 1.18$ and, for low values of the half-lives, there seems to be a power law behaviour with exponent $\hat{\alpha}_L = 0.88$. One can realize that $\hat{\alpha}_H + \hat{\alpha}_L = 2.06 \approx 2$. This behaviour has also been observed in the band transport of charges in amorphous semiconductors, for instance [13]. It would be interesting to explore a possible mathematical connection between both phenomena.

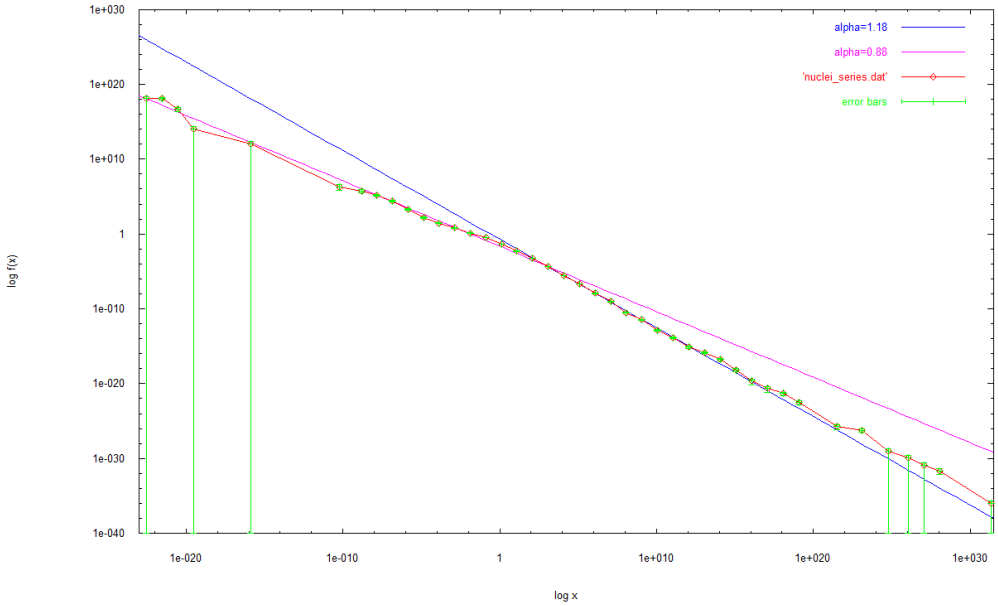


Figure 10: This figure shows, in purple, the fit with minimum $\hat{d}_{KS}$ for the power law at low half-lives ($\hat{\alpha}_L = 0.88$) and, in blue, the fit with minimum $\hat{d}_{KS}$ for the power law at high values of half-lives ($\hat{\alpha} = 1.18$).

6.3 Log-normal

The results for the log-normal distribution, as we mentioned, have been obtained with the same algorithm as the power law one. In this case, we have taken the lower bound from $3.16 \cdot 10^{-7}$ to $3.16 \cdot 10^{-1}$ seconds and the upper bound from $3.16 \cdot 10^{0}$ to $3.16 \cdot 10^{11}$ seconds. The intervals have been generated with the formula $a_i = 2\sqrt{5} \cdot 10^i$ where $i = -7, \ldots, -1$ for the increment of a, and the formula $b_j = 2\sqrt{5} \cdot 10^j$ where $j = 0, \ldots, 11$ for the increment of b. In table 5 we present the results for $a = 3.16 \cdot 10^{-2}$ seconds and for all possible values of b, since it is the range we have found the minimum of $\hat{d}_{KS}$.

b(s)	$\hat{\mu} \pm u(\hat{\mu})$	$\hat{\sigma}^2 \pm u(\hat{\sigma}^2)$	$\hat{d}_{KS}$	N	p
$3.16{\cdot}10^{0}$	0.51±0.02	9.23 ±0.51	0.019	664	0.97
$3.16{\cdot}10^{1}$	1.87±0.06	16.70±0.71	0.015	1106	0.97
$3.16{\cdot}10^{2}$	4.02±0.10	34.19±1.21	0.015	1595	0.87
$3.16{\cdot}10^{3}$	3.50±0.08	28.90±0.92	0.012	1977	0.92
$3.16{\cdot}10^{4}$	3.29±0.07	25.48±0.77	0.015	2212	0.70
$3.16{\cdot}10^{5}$	3.44±0.07	30.11±0.87	**0.010**	2393	**0.96**
$3.16{\cdot}10^{6}$	3.45±0.07	32.26±0.91	0.013	2491	0.75
$3.16{\cdot}10^{7}$	3.41±0.07	37.99±1.06	0.022	2571	0.18
$3.16{\cdot}10^{8}$	3.40±0.07	38.56±1.07	0.022	2598	0.14
$3.16{\cdot}10^{9}$	3.26±0.06	42.93±1.18	0.027	2626	0.04
$3.16{\cdot}10^{10}$	3.17±0.06	45.16±1.24	0.030	2638	0.02
$3.16{\cdot}10^{11}$	2.95±0.06	49.86±1.37	0.034	2651	0.01

Table 5: Results for the log-normal adjustment and the computed p-value, for a fixed lower bound of $a = 3.16{\cdot}10^{-2}$ seconds. Both $u(\hat{\mu})$ and $u(\hat{\sigma}^2)$ are lower levels for the statistical error associated to the paramters—see [14]. The smallest KS distance and corresponding p-value are in bold.

One can see in table 5 that the smallest KS distance ($\hat{d}_{KS} = 0.010$) is obtained for $a = 3.16 \cdot 10^{-2}$ and $b = 3.16 \cdot 10^{5}$ seconds, where the computed p-value takes the rate $p = 96\%$. Then, we accept as a best fit for our data, in the range delimited for these upper and lower bounds, a log-normal distribution with characteristic parameters $\hat{\mu} = 3.44$ and $\hat{\sigma}^2 = 30.11$. However, as one can see, there are several high p-values in the top of the table, then we can say that the null hypothesis must be accepted in these cases too. In figure 12 we show the fit with the above mentioned parameters and in figure 11 an evolution of $\hat{d}_{KS}$ as a function of a and b is presented.

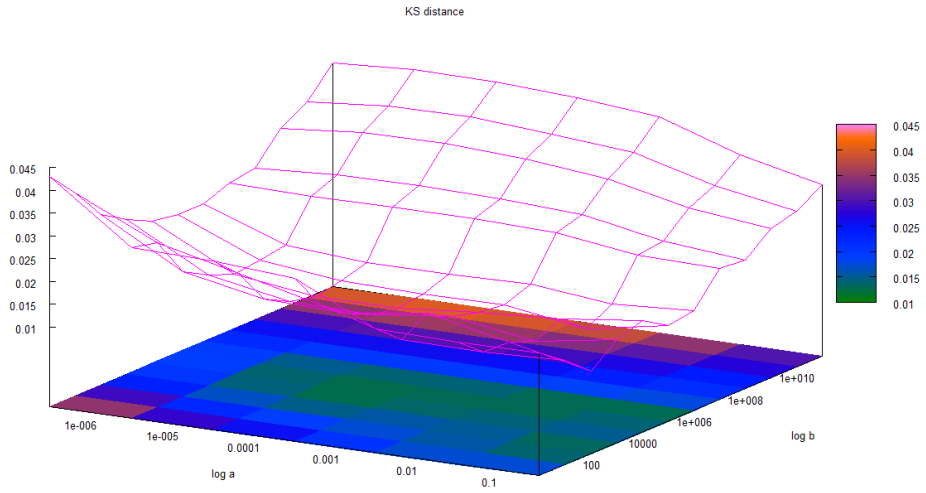


Figure 11: Surface described by the KS statistic, as a function of $\ln a$ and $\ln b$, in the case of the log-normal adjustment.

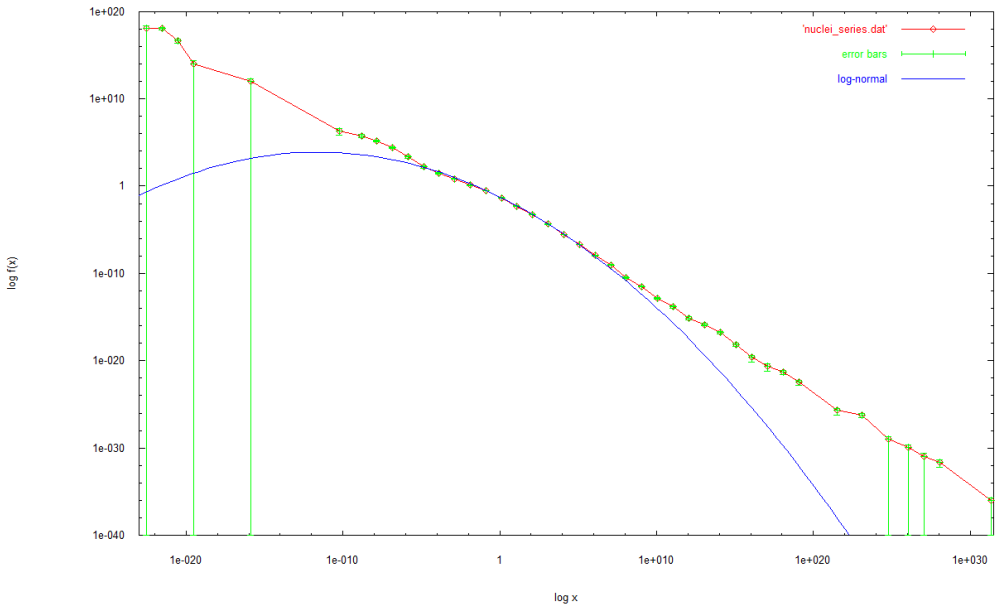


Figure 12: This is the fit of the log-normal density with parameters $\hat{\mu} = 3.44$ and $\hat{\sigma}^2 = 30.11$, with lower bound $a = 3.16 \cdot 10^{-2}$ s and upper bound $b = 3.16 \cdot 10^{5}$ s, where we have encountered the minimum $\hat{d}_{KS}$ and one of the most high p-values.

It should be noted that the number of data points N, in the fits commented above, is large. Specifically, the fit that gives the minimal $\hat{d}_{KS}$ has $N = 2393$ of the 3002 total data points. This makes the log-normal distribution a very representative one for most of the radioactive isotopes. To conclude the section, table 6 summarizes the main results obtained.

$a(s)$	$b(s)$	distribution	parameters
$2.61 \cdot 10^{-9}$	$1.47 \cdot 10^{-6}$	power law	$\hat{\alpha} = 0.88 \pm 0.08$
$3.16 \cdot 10^{-2}$	$3.16 \cdot 10^{5}$	log-normal	$\hat{\mu} = 3.44 \pm 0.07$, $\hat{\sigma}^2 = 30.11 \pm 0.87$
$1.47 \cdot 10^{2}$	$1.47 \cdot 10^{14}$	power law	$\hat{\alpha} = 1.18 \pm 0.01$
$1.47 \cdot 10^{9}$	-	pure power l.	$\hat{\alpha} = 1.08 \pm 0.01$

Table 6: This table shows a summary of the results found. The first three cases have the property of minimize the $\hat{d}_{KS}$ with an acceptable p-value. Concretely, in the cases of the truncated power law, the p-values for the results shown in the table are the highest ones. The last row shows the result with the highest p-value, for the analysis of the data, assuming a pure power law, which does not coincide with the minimal $\hat{d}_{KS}$.

7 Conclusions

In this work, the method proposed by Clauset *et al.* in [1] to fit power laws distributions and to test the goodness of fit by means of Monte-Carlo simulation has been implemented and generalized to the case of truncated power laws and log-normal distributions. Application of the methods to the half-lifes of radioactive elements reveals both the interest of using such record as an illustration of a data set with a broad range of variability and a long-tailed behaviour, and also the relative limitation of the optimization method proposed by Clauset *et al.* We see that a "blind" application of the recipe is not recommended at all.

We can say that the method and criteria used bring to accept that there is a power law behaviour of the nuclide data in the range from $1.47 \cdot 10^2$ to $1.47 \cdot 10^{14}$ seconds with an exponent of $\alpha = 1.18$. We speculate that the slight deviation to an exponent $\alpha = 1.08$ for the highest half-lives could be due to an artifact caused by a biased sampling for these values. Anyhow, a power law tail implies the lack of a characteristic scale in the half-lifes, in this case [15]. In addition, at low values of half-lives the data seems to have a power law behaviour with approximated exponent $\alpha = 0.88$. Nevertheless, due to the few number of data at low half-lives, we can not be certain at all that the data in this range is drawn from a power law distribution.

Moreover, we have encountered a clear log-normal behaviour of the data in the intermediate values of half-lives, between the two power laws, where the number of data points is very large.

It would be interesting to determine with a certain precision for which value of the half-life the transition from the log-normal distribution to the power law distribution takes place. A maximum likelihood ratio test could provide the necessary information. Preliminary analysis shows, however, that it does not seem easy to find such value, as the ratio is very close to one for the range in which both distributions overlap, so the differences between both distributions would not be significant.

As a future objective, we could look for a unique probability distribution that could account for the different behaviours of the half-life distribution. Different limit behaviours should yield a power law or a log-normal, depending on the value of the half-life. Certainly, in this case, we would need to modify our estimation procedure, as it would not be possible to find a closed expression for the normalization constant, which should be determined numerically for each value of the parameters.

Finally, the statitical findings reported here open interesting physical ques-

tions. Which is the reason that the radioactive nuclides obey such a clear probability law (first power law, then lognormal, and then a somewhat steeper power law)? How are these findings related to the different modes of radioactive decay? Does the distribution change when it is conditioned to alpha emission, beta decay, etc.?

In summary, our work offers new opportunities of research, which can unveil some unknown properties of radioactive nuclei thanks to careful statistical analysis.

8 Acknowledgments

I would like to express my gratitude to my supervisors, Álvaro Corral and Ramon Nonell, whose expertise, understanding and patience, have made possible the achievement of this research work. Moreover, I gratefully acknowledge the financial support of the Centre de Recerca Matemàtica and of the CERMET project from the Facultat de Matemàtiques i Estadística of the UPC.

9 Appendix

9.1 Cumulative distribution function

Since the cumulative distribution function of a given continuous random variable X with probability density $f(x)$ is defined as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x')dx', \tag{25}$$

their complementary is defined as

$$F^c(x) = P(X > x) = \int_{x}^{\infty} f(x')dx' \tag{26}$$

or, equivalently, $F^c(x) = 1 - F(x)$.

The routine we have used to estimate the d_{KS} between two distributions needs the formula of the CDF, or their complementary, of the distribution as an input. For simplicity, in the cases of power laws it has been useful to implement the complementary function and in the case of log-normal we have implemented the cumulative function. Here we show the calculations:

- Complementary CDF of a power law in $[a, \infty)$:

$$\begin{aligned} F^c(x) &= C \int_x^\infty \frac{1}{x'^\alpha}\, dx' \\ &= \frac{C}{1-\alpha} \left[x'^{1-\alpha}\right]_x^\infty \\ &= \frac{C}{1-\alpha} x^{1-\alpha} \\ &= \left(\frac{x}{a}\right)^{1-\alpha} \end{aligned}$$

  where C comes from the formula (3).

- Complementary CDF of a power law in $[a, b]$:

$$\begin{aligned} F^c(x) &= C \int_x^b \frac{1}{x'^\alpha}\, dx' \\ &= \frac{C}{1-\alpha} \left[x'^{1-\alpha}\right]_x^b \\ &= \frac{C}{1-\alpha} \left[b^{1-\alpha} - x^{1-\alpha}\right] \\ &= \frac{1 - \left(\frac{x}{b}\right)^{1-\alpha}}{1 - \left(\frac{a}{b}\right)^{1-\alpha}}. \end{aligned}$$

where C comes from the formula (5).

- Using the notation employed in subsection 4.3, the complementary CDF of a normal distribution in the interval $[A, B]$ is the following:

$$\begin{aligned} F^c(x) &= C\int_x^B e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt' \\ &= C\sqrt{2\sigma^2}\int_{\frac{x-\mu}{\sqrt{2\sigma^2}}}^{\frac{B-\mu}{\sqrt{2\sigma^2}}} e^{-z^2}\,dz \\ &= C\sqrt{2\sigma^2}\left(\int_0^{\frac{B-\mu}{\sqrt{2\sigma^2}}} e^{-z^2}\,dz - \int_0^{\frac{x-\mu}{\sqrt{2\sigma^2}}} e^{-z^2}\,dz\right) \\ &= C\frac{\sqrt{2\pi\sigma^2}}{2}\left[\mathrm{erf}\left(\frac{B-\mu}{\sqrt{2\sigma^2}}\right) - \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right] \\ &= \frac{\mathrm{erf}\left(\frac{B-\mu}{\sqrt{2\sigma^2}}\right) - \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)}{\mathrm{erf}\left(\frac{B-\mu}{\sqrt{2\sigma^2}}\right) - \mathrm{erf}\left(\frac{A-\mu}{\sqrt{2\sigma^2}}\right)} \end{aligned}$$

where C comes from the formula (8). In the second equality we have applied the standard change of variable $z = (t-\mu)/\sqrt{2\sigma^2}$.

9.2 Computation of $S_N(x)$

The computation of $S_N(x)$ is very simple. If the N events are located at values x_i, $i = 1, \ldots, N$, then $S_N(x)$ is the function giving the fraction of data points to the left of a given value x. Then, previously sorting the data $x_1, \ldots, x_N$ in ascending order, the value of the function at given i will be $S_N(x_i) = i/N$. Since we have worked with the complementary of the CDF, the corresponding complementary of $S_N(x)$ is $S_N^c(x) = 1 - S_N(x)$ and, obviously, $S_N(x_i) = 1 - i/N$.

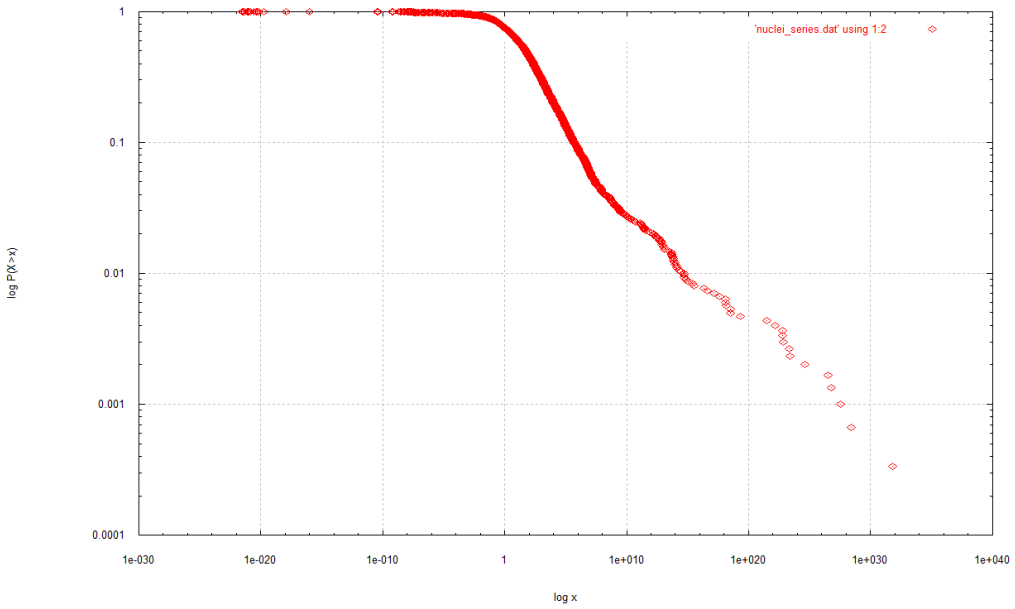


Figure 13: This plot shows the empirical complementary CDF, $S_N^c(x)$, for our half-lives data.

9.3 The Transformation Method

Proposition 9.1. *Let X a random variable with cumulative distribution function $F_X(x)$. Define a new random variable $Y = F_X(x)$, then Y is uniformly distributed $Y \sim unif[0,1]$.*

Proof. Computing the cumulative distribution function for Y occurs the following

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P[F_X(x) \leq y] \\ &= P[X \leq F_X^{-1}(y)] \\ &= F_X(F_X^{-1}(y)) \\ &= y \end{aligned}$$

so $Y \sim \text{unif}[0,1]$. □

References

[1] A. Clauset, C. R. Shalizi and M. E. J. Newman, *Power-law distributions in empirical data.* SIAMR **Vol. 51** (2009), 661–703.

[2] M. E. J. Newman, *Power laws, Pareto distributions and Zipf's law.* Contemporary Physics **Vol. 46** (2005), 323–351.

[3] A. Corral, *Statistical Tests for Scaling in the Inter-Event Times of Earthquakes in California.* International Journal of Modern Physics B **Vol. 23** (2009), 5570–5582.

[4] E.P. White, B.J. Enquist, J.L. Green, *On estimating the exponent of power-law frequency distributions.* Ecology, **Vol. 89** (2008), 905–912.

[5] S. Hergarten, *Self-Organized Criticality in Earth Systems.* Springer (2002).

[6] P. Marcillac, N. Coron, G. Dambier, J. Leblanc, J. Moalic, *Experimental detection of α-particles from the radioactive decay of natural bismuth.* Nature, **Vol. 422** (2003), 876–878.

[7] A. Corral, *Dependence of earthquake recurrence times and independence of magnitudes on seismicity history.* Tectonophysics **424** (2006), 177–193.

[8] J.A. Rice, *Mathematical Statistics and Data Analysis.* Wadsworth & Brooks/Cole Advanced Books & Software (1988).

[9] V.K. Rohatgi, A.K.Md. Ehsanes Saleh, *An Introduction to Probability and Statistics.* Wiley Series in Probability and Statistics (2001), Second Edition.

[10] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in Fortran 77.* Cambridge University Press (1992), Second Edition.

[11] S.M. Ross, *Introduction to Probability Models.* Academic Press (2007), 9th edition.

[12] I.B. Aban, M.M. Meerschaert, and A.K. Panorska, *Parameter Estimation for the truncated Pareto Distribution.* Journal of the American Statistical Association, **Vol. 101** No.473 (2006), 270–277.

[13] H. Scher, M.F. Shlesinger, J.T. Bendler, *Time-scale invariance in transport and relaxation.* Physics Today, **Vol. 44** (1991), 26–34.

[14] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans.* SIAM, **Vol. 51** No. 4 (1982), 661–703.

[15] A. Corral, *Física estadística dels fenòmens catastròfics.* Fundació Caixa Sabadell (2007), p. 45-67.

Campus de Bellaterra, Edifici C
08193 Bellaterra, Spain
Tel.: +34 93 581 10 81
Fax: +34 93 581 22 02
crm@crm.cat
www.crm.cat