Neutrino mass and Extreme Value Distributions in $\beta$-decay

J. G. Esteve$^*$ and Fernando Falceto$^\dagger$
Departamento de Física Teórica, Universidad de Zaragoza, 50009 Zaragoza, Spain and
Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), 50009 Zaragoza, Spain
(Dated: January 9, 2014)

We propose a new application of the Extreme Value Theory for distributions with compact support. The novelty of our proposal is the use of these tools to estimate the neutrino mass from the energy spectrum of electrons in $\beta$-decay. In this way the dependence of the result on the mass of the neutrino is considerably enhanced increasing the sensitivity of the experiment.

PACS numbers: 05.10.Cc

I. INTRODUCTION

Neutrino masses play an important role in different areas of Physics ranging from Nuclear to High Energy Physics or Astrophysics. Up to date, however, we only know upper bounds to their values and their determination remains an outstanding problem in Physics. One of the most promising procedures to unravel the mass of the electronic (anti)neutrino is to study the electronic energy spectrum in the $\beta$-decay, where the maximum energy allowed to the electron depends on the mass of the (anti)neutrinos (in what follows we will not differentiate between the masses of neutrino and antineutrino, as we assume an exact CPT symmetry). However, the ratio of the electronic neutrino mass $m_\nu$ to the electron mass $m_e$ is very small, $m_\nu/m_e < 10^{-5}$ ([1] - [3]), and the probability of emission of the electron depends at least quadratically on this quotient. This implies that the spectrum of the emitted electrons is very little affected by the value of $m_\nu$. Only near the end point of the spectrum, where the energy of the emitted antineutrino is close to $m_\nu c^2$, there is a stronger dependence in the neutrino mass as the highest energy of the electron varies linearly with $m_\nu c^2$. This adds new difficulties to measure the neutrino mass in $\beta$-decay, since one must explore the region very near the end point, where the occurrence of an event has very small probability.

For these reasons, it is important to find new strategies to analyze the data with higher sensitivity in this part of the electron spectrum. This goal can be achieved with the help of the Extreme Value Theory (EVT) for random variables of compact support [4], as it is described below.

EVT studies the probability distribution of the maximum of $n$ independent, identically distributed random variables ([3] - [18]). Its interest in connection to the problem of determining the neutrino mass is clear if one considers that, as mentioned before, the influence of the neutrino mass on the energy distribution of the electronic emission, is maximal at the upper bound. Then, if we focus on the electrons emitted with the highest energy among $n$ emissions we should expect a deeper insight of the region near the upper bound of the spectrum. In the next sections we will introduce the main tools in EVT and we will present quantitative results for the Tritium $\beta$-decay showing the enhancement in the sensitivity of the experiment.

II. EXTREME STATISTICS

Let us consider a normalized probability density $\rho(x)$ with support in $[0, 1]$, i.e. $\rho(x) = 0$ if $x \notin [0, 1]$, and $\mu(x)$ its cumulative distribution function:

$$\mu(x) = \int_0^x \rho(s)ds.$$  

The distribution function of the maximum of $n$ independent identically distributed (i.i.d.) random variables with probability density $\rho(x)$ is then:

$$\mu_n(x) = \mu(x)^n,$$  

(1)

$^*$ Corresponding author.
$^\dagger$Electronic address: esteve@unizar.es
$^\ddagger$Electronic address: falceto@unizar.es
as it can be easily argued. In fact, note that $\mu(x)$ represents the probability of the random variable taking a value smaller than $x$. Therefore, the probability of $n$ i.i.d random variable taking all of them a value smaller than $x$ is the $n$th power of the former. Hence, the probability density is:

$$
\rho_n(x) = \frac{d}{dx}\mu_n(x) = n\mu(x)^{n-1}\rho(x).
$$

We are interested in the large $n$ limit where it is evident that $\rho_n(x)$ squeezes at the end point $x = 1$ independently of the initial probability density (provided $x = 1$ is in the support of $\rho$). Therefore, we lose any sign of the initial distribution.

A way to overcome this situation is to renormalize, at every step, the random variable. As it is shown in [4] a convenient renormalization is to transform the variable $x$ into $x' = x^{n^{-\beta}}$. This transformation has the virtue of preserving the support $[0, 1]$, which means that the new random variable $x'$ takes the same range of values as the original one. Also, if we chose the appropriate $\beta$, the limit of infinite $n$ provides a non trivial probability distribution.

The problem of choosing the appropriate value for $\beta$ was addressed in [4] using renormalization group techniques. There it is shown that $\beta$ depends solely on the behavior of $\rho$ in the vicinity of $x = 1$. More specifically, assuming that

$$
\rho(x) \sim \gamma(1-x)^{\alpha-1},
$$

(where $\alpha > 0$ in order to have a normalizable density) the appropriate transformation is implemented with $\beta = 1/\alpha$.

Once the right transformation has been chosen one can safely take the large $n$ limit and obtain the Renormalization Group fixed point

$$
M(x) = \lim_{n \to \infty} \mu_n(x') = e^{-\lambda(-\log(x))\rho},
$$

where $x' = x^{n^{-1/\alpha}}$ and $\lambda = \gamma/\alpha$. Observe that the limiting probability distribution $M$ depends only on the two parameters $\alpha$ and $\gamma$ that determine the behavior of $\rho$ at the maximum of its support. All other details of the initial distribution are swept out. If we had considered a different rescaling ($\beta \neq 1/\alpha$) then the limiting probability density would have been trivial in the sense that it accumulates either at 0 or at 1. In fact, in this case we have

$$
\lim_{n \to \infty} \mu_n(x'^{n^{-\beta}}) = H(x - x_0),
$$

with $H$ being the Heaviside step function and $x_0 = 0$ for $\beta > 1/\alpha$ and $x_0 = 1$ for $\beta < 1/\alpha$.

These properties are key to our purposes, due to the fact that the energy spectrum of the electron emitted in the $\beta$ decay changes drastically at the upper limit depending whether the neutrino is massless or not. Namely, if the neutrino were massless the probability density for the electronic energy would behave like $(E_{\text{max}} - E)^2$ ($\alpha = 3$). On the other hand, if it were massive the behavior would be $(E_{\text{max}} - E)^{1/2}$ ($\alpha = 3/2$). Therefore we expect that the sensitivity of the experiment to the neutrino mass should be very much enhanced when the results are analyzed with the tools provided by EVT.

### III. APPLICATION TO TRITIUM $\beta$-DECAY

Now we are going to apply these ideas to the problem of determining with more precision the neutrino mass in $\beta$-decay experiments. Although the method is general, we will focus on the $\beta$-decay of Tritium. This phenomenon is specially suited for the purpose of determining the neutrino mass for two reasons. First, it is a super allowed decay (12.3 years of half-live) and, consequently, the nuclear matrix element does not depend on the energy of the emitted electron. Second, it has a relatively low end point energy, which increases the sensitivity of the experiment to the neutrino mass (see [15] for a review on the neutrino mass limit from the Tritium beta decay).

The probability of emission of an electron, with total energy $E$ ranging from $m_c^2$ to $E_0 - m_c^2$, is given by

$$
p_e(E) = N_e G(E) (E_0 - E) [(E_0 - E)^2 - m_e^2 c^2]^{1/2},
$$

where $N_e$ is a normalization constant, $E_0$ the total available energy, i. e. $E_0 = \epsilon_0 + m_c^2$ with $\epsilon_0$ the end point energy, and

$$
G(E) = \frac{E^2}{1 - \epsilon_0 - \sqrt{E^2 - m_e^2 c^4}}.
$$
is related to the Fermi function which accounts for the interaction with the nuclear Coulomb field. In the case of atomic Tritium we shall take $\epsilon_0 = 18.560 \text{ KeV}$ \[12\].

Strictly speaking, $m_\nu^2$ is the squared mass of the electron antineutrino, which should be expressed in terms of the weighted average of the mass eigenstates $m_i$ ($i = 1, 2, 3$) as $m_\nu^2 = \sum_{i=1,2} |U_{ei}|^2 m_i^2$ being $|U_{ei}|^2$ the weights known from the neutrino oscillation experiments. For the sake of clarity we have neglected in \[10\] the sum over final states. Which does not affect the conclusions while it can be implemented numerically.

To simplify the analysis it will be convenient to introduce dimensionless, rescaled variables for the electron kinetic energy and the neutrino mass. Namely, define

$$x = \frac{E - m_e c^2}{E_0 - m_e c^2}, \quad y = \frac{m_\nu c^2}{E_0 - m_e c^2},$$

and the probability density in terms of $x$ is

$$\rho(x) = N F(x) (1 - x) [(1 - x)^2 - y^2]^{1/2}$$

where $0 < x < 1 - y$, $N$ is the new normalization constant and $F(x) = G(x(E_0 - m_e c^2) + m_\nu c^2)$. Now by monitoring the energy of the electrons in the $\beta$-decay of Tritium we could estimate the probability density $\rho$ and deduce the value of $y$. The problem is that, as we mentioned before, the dominant correction to the limit of zero neutrino mass goes like $y^2 \frac{21}{24}$ and, given the small value for $y \sim 10^{-5}$, the corrections due to the mass of the neutrino are very small. As we shall see, the EVT for $n$ events provides an enhancement of the correction by a factor $n^{2/3}$ making it easier to detect.

The experimental set up should focus on the electron of highest energy among $n$ decays. Repeating this process many times, one obtains the probability of occurrence of such an extreme value for the electron energy. Note that the number of decays we consider at each step can be determined if we know how many atoms of Tritium we have in.

The cumulative distribution for the highest energy events is $\mu_n(x) = \mu(x)^n$ where, as before, $\mu$ is the cumulative distribution function for the probability density in \[8\]. In order to get a non trivial limit (in the case of zero neutrino mass) we shall rescale the variable and define $M_n(x) = \mu_n(x^{n^{-1/3}})$. To estimate the size of the corrections due to the non-zero neutrino mass we first expand $\rho$ in powers of $y$,

$$\rho(x) = a(x)(1 - x)^2 + y^2 b(x) + O(y^3)$$

where

$$a(x) = Q^{-1} F(x),$$
$$b(x) = [(1 - x)^2 Q^{-1} R - 1] Q^{-1} F(x) \frac{2}{2},$$

with $Q = \int_0^1 F(x) (1 - x)^2 dx$ and $R = \int_0^1 F(x) dx$. Then the cumulative distribution $\mu$ can be similarly written

$$\mu(x) = 1 - A(x) (1 - x)^3 - y^2 B(x) (1 - x) + O(y^3).$$

For the moment we do not specify the concrete form of $A(x)$ and $B(x)$.

Now we can easily expand the logarithm of $M_n(x) = \mu(x^{n^{-1/3}})^n$:

$$\log M_n(x) = A(x^{n^{-1/3}}) \log^3 x + n^{2/3} y^2 B(x^{n^{-1/3}}) \log^3 x + \ldots$$

where the dots represent subdominant contributions for large $n$ and small $y$ and we have used $1 - x^{n^{-1/3}} = -n^{-1/3} \log x + \ldots$. Therefore, taking the large $n$ limit (with $n^{1/3} y < 1$) in the arguments of $A$ and $B$ and exponentiating the previous expression we obtain

$$M_n(x) = e^{\lambda \log^3 x (1 - \frac{3}{2} n^{2/3} y^2 \lambda \log x + \ldots)}.$$  \(9\)

Where $\lambda = A(1) = (1/6) \mu''(0)|_{y=0}$ and, given the relation between $\mu$ and $\rho$, we have $\lambda = (1/6) \rho''(0)|_{y=0} = a(1)/3 \approx 2.04208.$
The expansion for the probability density is obtained by taking the derivative of (9) and it reads

\[ P_n(x) = P(x) \left( 1 - \frac{1}{2} n^{2/3} y^2 (3\lambda \log x + \log^2 x) + \ldots \right), \]

where \( P(x) \) is the probability density in the limit of zero neutrino mass, i.e.

\[ P(x) = 3\lambda \frac{\log^2 x}{x} e^{\lambda \log^3 x}. \]

The expressions above are valid provided \( n \gg 1 \) with \( n^{1/3} y < 1 \) which is the real situation, given the actual upper bounds for \( m_\nu \) and the accessible experimental values for \( n \) (of the order of \( 10^{12} \), as we will discuss below). If we had considered the large \( n \) limit in a more strict sense, \( n \gg 1 \) and \( n^{1/3} y \gg 1 \), the results would have changed and the difference between the massless and massive case would have been more dramatic. In fact, after performing the appropriate rescaling, we would have obtained for very large \( n \):

\[ \log M_n(x) \propto \begin{cases} -|\log x|^3 & (m_\nu = 0) \\ -|\log x|^{3/2} & (m_\nu \neq 0). \end{cases} \]

However, as we argued before, in the present situation this limit is not attainable. Therefore, we shall restrict ourselves to the case when \( n^{1/3} y < 1 \) using the expression (9) for \( M_n(x) \).

In our analysis of the Tritium \( \beta \)-decay we address the possibility of measuring a neutrino mass of the order \( m_\nu/m_e = 10^{-6} \). In Fig. 1 we show a comparison between the exact values of \( M_n(x) \) for \( y = 0 \) and the corresponding value of the rescaled variable \( y = 2.8 \times 10^{-5} \) for non-zero neutrino mass. We also show the curve obtained using the perturbative expansion (9) with \( n^{2/3} y^2 = 0.078 \). We can see there that the perturbative expansion fits well with the exact values \( (y \neq 0) \) showing that the first correction to the integrated distribution for \( y = 0 \) is proportional to \( n^{2/3} y^2 \).

Next, in Fig. 2 we plot the cumulative distribution \( M_n(x) \) for \( n = 10^{12} \), \( m_\nu = 0 \) and \( m_\nu c^2 = 0.5 \text{ eV} \). We choose this value of \( n \) for the number of disintegrations because as the figure shows, the distributions corresponding to \( m_\nu c^2 = 0.5 \text{ eV} \) and \( m_\nu = 0 \) are clearly separated. Furthermore, it is experimentally achievable with the present available devices. For example the decay rate planned for future experiments is about \( 10^{11} \text{ Bq} \), so one could record an extreme event every 10 seconds.

In Fig. 3 we compare the probability distribution \( P_n(x) \) for \( n = 10^{12} \) and different values of \( m_\nu \). The graph illustrates how the bound on the neutrino mass, for this \( n \), can be lowered by one order of magnitude with respect to its present value. However, it would be very difficult to go one order of magnitude further, as it would imply to increase \( n \) by a factor 10^3.

We end our discussion by studying the expectation value of the renormalized variable \( x^{n^{1/3}} \), that is also sensitive to the neutrino mass. In this case we also discuss how the precision in the measurement of the electronic energy affects
FIG. 2: $M_n(x)$ for $n = 10^{12}$ and $m_{\nu}/m_e = 10^{-6}$ (continuous line), and for $m_{\nu} = 0$ (discontinuous line). The vertical line indicates the value of $(1 - y)^{n^{-1/3}}$.

FIG. 3: $P_n(x)$ for $n = 10^{12}$ and $m_{\nu}c^2 = 1$ eV (thinnest line), $m_{\nu}c^2 = 0.5$ eV (continuous thicker line), $m_{\nu}c^2 = 0.25$ eV (doted line) and $m_{\nu} = 0$ (discontinuous line).

the determination of the neutrino mass. We have:

$$\langle x^{n^{1/3}} \rangle = \int_0^{1-y} x^{n^{1/3}} \frac{d}{dx} \mu_n(x) dx$$

$$= \langle x^{n^{1/3}} \rangle_{y=0} - \frac{3}{2} n^{2/3} y^2 \lambda \int_0^{\infty} e^{-\lambda t^2} t dt + \ldots,$$

where $\langle x^{n^{1/3}} \rangle_{y=0}$ is the mean value for $m_{\nu} = 0$. In our case

$$\langle x^{n^{1/3}} \rangle \approx 0.510864 - 0.496771 n^{2/3} y^2 + \ldots, \quad (11)$$

where the dots mean corrections of order $n^{-1/3}$ and $n^{4/3} y^4$.

We assume now an indetermination $\Delta E$ in the measurement of the electronic energy due to the resolution of the experimental devices. This is related to the indetermination of the neutrino mass ($y$ in the adimensional variables) by the following expression,

$$\Delta E = 0.993542 \times (E_0 - m_e c^2) \frac{n^{1/3} y^2 \Delta y}{\langle x^{n^{1/3}} \rangle y}$$

which one easily gets from (11).

Therefore, assuming $m_{\nu} \sim 0.5$eV, we can determine it within 50% of relative precision, provided the energy resolution is $\Delta E \sim 0.88$eV and $n = 2 \times 10^{13}$ (an account every 200 seconds in an experiment with $10^{11}$ Bq). This estimates are close to the values that are expected from future experiments.
IV. CONCLUSIONS

In this paper, we briefly presented the main tools of Extreme Value Statistics for distributions of compact support. A subject that can be useful for analyzing the electron spectrum in the $\beta$-decay of tritium, in order to determine an upper limit for the electronic neutrino. Actually, the main difference in the spectrum with a massive or a massless neutrino is its behavior at the tail of the distribution, a region that can be deeply explored thanks to the extreme value theory.

We have performed an analytic study of the expectation variable of the renormalized energy of the fastest emitted electron among $n$ disintegrations. We have shown that the expectation variable depends linearly on $n^{2/3}y^2$ and we have determined the coefficients. In this way we show that the dependence on the square mass of the neutrino (in dimensionless units) $y^2$ is enhanced by the factor $n^{2/3}$, which increases the sensitivity of the experiment. For instance, in the present experimental facilities, we can reach $n \approx 10^{12}$, which means that neutrino mass upper bounds of the order of 0.5eV could be obtained. In conclusion, we believe that extreme value theory can be useful to analyze the electronic energy spectrum in $\beta$-decay and it is worth implementing in present and future experiments.

Acknowledgments: Research partially supported by grants 2012-E24/2, DGIID-DGA and FPA2012-35453, ENE2009-07247, Ministerio de Ciencia e Innovación (Spain).

[1] J. Beringer et al. (Particle Data Group), Phys. Rev. D86, 010001 (2012).
[2] Kraus Ch. et al., Eur. Phys. J. C 40, 447, (2005).
[3] Lobashev V. M. et al., Phys. Lett. B 460, 227, (1999).
[4] I. Calvo, J. C. Cuchí, J. G. Esteve and F. Falceto, Phys. Rev. E 86, 041109 (2012).
[5] L. de Haan and A. Ferreira, Extreme Value Theory: An Introduction (Springer, 2006).
[6] W. Weibull, J. Appl. Mech.-Trans. ASME 18, 293 (1951).
[7] G. Jona–Lasinio, Nuovo Cimento 26, 98 (1975).
[8] I. Calvo, J. C. Cuchí, J. G. Esteve and F. Falceto, J. Stat. Phys. 141, 409 (2010).
[9] G. Gyorgyi, N. R. Moloney, K. Ozogany and Z. Racz, Phys. Rev. Lett. 100, 210601 (2008).
[10] G. Gyorgyi, N. R Moloney, K. Ozogany, Z. Racz and M. Droz, Phys. Rev. E 81, 041135 (2010).
[11] E. Bertin and G. Gyorgyi, J. Stat. Mech., P08022 (2010).
[12] M. Fréchet, Ann Soc. Polonaise Math. 6, 93 (1927).
[13] M.R. Fisher and L.H.C. Tippett, Proc. Cambridge Philos. Soc. 24, 180 (1928).
[14] B. V. Gnedenko, Ann. Math. 44, 423 (1943).
[15] R.L. Smith, Breakthroughs in statistics, (S. Kotz and N. L. Johnson eds.) Springer-Verlag, (1993).
[16] E. Pancheva, ProbStat Forum 3, 11 (2010).
[17] E. Pancheva, Lecture Notes in Math., 1155, 284 (1984).
[18] N.R. Mohan and S. Ravi, Theory Probab. Appl. 37, 632 (1991).
[19] E. W. Otten and C. Weinheimer, Rep. Prog. Phys. 71, 086201 (2008).
[20] Note that, although apparently $\rho$ depends quadratically on $y$, there could have a linear contribution from the normalization constant. However, one can deduce that it is absent due to the fact that $\rho$ vanishes in the upper limit $x = 1 - y$. 