Super-Kamiokande NOTE 98-03

Confidence Limit Calculation for Setting Proton Decay Lifetime Limits

B. Viren
State University of New York, Stony Brook

May 22, 1998

Abstract

Making use of Bayes' theorem, the method used for calculating proton decay lifetime limits is detailed, with particular emphasis on the prior probability distribution for the background. The method is then applied to various example cases.

1 Introduction

When one observes N particles for a time T and finds a nonzero number of decays, n, the rate of decay is estimated simply by,

$$\Gamma = \frac{n}{NT\epsilon}, \tag{1}$$

where ϵ is the detection efficiency to observe this mode's decay products. However, when n is zero, a rate can not be accurately calculated. This is because n is no longer a good estimator of the true mean. Naively one could calculate a decay rate of zero, however, this may be wrong for it may be the case that if we had observed more particles for a longer time a nonzero number of decays would have been seen. So, instead, a limit on the decay rate at some confidence level is set.

In simple situations, it is common to use the *Poisson Process with Background* method [1]. This simple Poisson limit is set by finding a value, μ_{limit}, which is the mean of a Poisson distribution such that there would be a certain probability, usually 90%, that any value chosen from this distribution would be larger than μ_{limit}. To find the limit on the decay rate, μ_{limit} is then substituted for n in (1).

However, what this method lacks is the ability to take into consideration any uncertainties we may have in the parameters that go in to the limit, namely, the estimated efficiency, ϵ, the estimated background, b, and the exposure, $\lambda = NT$. Because the concept of this limit is innately based on *confidence* we must somehow include the degree to which we are or are not confident in our estimated quantities.

Through the use of Bayes' theorem and the assignment of prior probability distributions (priors) it is possible to use the knowledge of our uncertainties to calculate a limit which better corresponds to the confidence we quote. The use and validity of Bayes' theorem will not be discussed or justified here as there are other sources [2, 3] that would do the topic better justice.

In this note, the method used for the first $p \rightarrow e^+\pi^0$ proton decay publication [4] will be detailed and this method will be applied to the primary, and secondary analysis that is presented in the publication. The calculation described in the next section is heavily based on the method outlined in Ref. [5].

2 The Method

The expected mean number of candidates, μ, in our data sample is,

$$\mu = \Gamma\lambda\epsilon + b. \tag{2}$$

Where, as above, Γ is the true decay rate (for a given decay mode), $\lambda = NT$ is the true exposure, ϵ is the true efficiency to detect the decay products, and b is the true mean number of background events in our sample. Because this is a counting experiment the probability to detect n events follows the Poisson distribution with Poisson parameter μ. That is, given the parameters: Γ, λ, ϵ and b, and all conditions of the experiment, I, the probability to detect n events is,

$$\mathbf{P}(n|\Gamma\lambda\epsilon bI) = \frac{e^{-(\Gamma\lambda\epsilon+b)}(\Gamma\lambda\epsilon + b)^n}{n!}. \tag{3}$$

Applying Bayes' theorem allows us to write,

$$\mathbf{P}(\Gamma\lambda\epsilon b|nI) = A\mathbf{P}(n|\Gamma\lambda\epsilon bI)\mathbf{P}(\Gamma\lambda\epsilon b|I). \tag{4}$$

The constant of proportionality, A, will be found by insisting that $\mathbf{P}(\Gamma\lambda\epsilon b|nI)$ is normalized. Given a particular set of cuts, the decay rate, efficiency, exposure, and background are all independent of one another. Therefore, the value $\mathbf{P}(\Gamma\lambda\epsilon b|I)$ can be separated into its constituents giving,

$$\mathbf{P}(\Gamma\epsilon\lambda b|nI) = A\mathbf{P}(n|\Gamma\epsilon\lambda bI)\mathbf{P}(\Gamma|I)\mathbf{P}(\epsilon|I)\mathbf{P}(\lambda|I)\mathbf{P}(b|I). \tag{5}$$

The quantities, $\mathbf{P}(\Gamma \mid I)$, $\mathbf{P}(\epsilon \mid I)$, $\mathbf{P}(\lambda \mid I)$ and $\mathbf{P}(b \mid I)$, are known as the prior probability distributions for Γ, ϵ, λ and b respectively. These distributions codify the state of our knowledge of each parameter before the outcome of our experiment is known. The important subject of the choice of priors is discussed in section 3. Given the priors, the joint probability distribution in (5) is marginalized with respect to the so called nuisance variables: ϵ, λ and b. This is done because we are not strictly interested in bettering our knowledge of these parameters. Rather we are interested only in what our experiment can tell us about the decay rate, the distribution of which we can now write as,

$$\mathbf{P}(\Gamma \,|\, nI) = \int\!\!\int\!\!\int \mathbf{P}(\Gamma\epsilon\lambda b \,|\, nI)\, d\epsilon\, d\lambda\, db. \tag{6}$$

At this point the normalization constant, A, can be resolved by demanding,

$$\int_0^\infty \mathbf{P}(\Gamma \,|\, nI)\, d\Gamma = 1. \tag{7}$$

Once it is known how the decay rate is distributed it is trivial to calculate a limit using a definition similar to that used to find μ_{limit} in the simple Poisson limit. Namely, we calculate a value, Γ_{limit}, such that there is some probability, equal to our confidence level (CL), that we would choose a value from $\mathbf{P}(\Gamma \,|\, nI)$ no higher than Γ_{limit}. This is done by solving,

$$\mathrm{CL} = \int_0^{\Gamma_{limit}} \mathbf{P}(\Gamma \,|\, nI) d\Gamma, \tag{8}$$

for Γ_{limit}.

2.1 Reduction to Simple Poisson Limit

The method just described will reduce to the simple Poisson limit [1] when systematic uncertainties go to zero and the prior for the decay rate is taken to be uniform. To show this, $\Gamma\epsilon\lambda$ is written simply as s the mean signal expected in the data. Following the above method, the mean number of events is,

$$\mu = s + b, \tag{9}$$

and the probability to observe n events given the mean signal, s and background, b, is,

$$\mathbf{P}(n \,|\, sbI) = \frac{e^{-(s+b)}(s+b)^n}{n!}. \tag{10}$$

Applying Bayes' theorem gives,

$$\mathbf{P}(sb \,|\, nI) = A\mathbf{P}(n \,|\, sbI)\mathbf{P}(s \,|\, I)\mathbf{P}(b \,|\, I). \tag{11}$$

In order to reduce to the simple Poisson limit, we must take the prior for the signal, $\mathbf{P}(s \mid I)$, to be uniform, and we must assume that the background is known with certainty, ie. $\mathbf{P}(b \,|\, I) = \delta(b - b_0)$. Marginalizing then gives,

$$\mathbf{P}(s \,|\, nI) = A \int_0^\infty \frac{e^{-(s+b)}(s+b)^n}{n!} \delta(b - b_0)\, db = A \frac{e^{-(s+b_0)}(s+b_0)^n}{n!}. \tag{12}$$

So to find the limit on the signal, s_{lim} at a confidence CL one must solve, after enforcing normalization and substituting into (8),

$$\mathrm{CL} = \frac{\int_0^{s_{lim}} e^{-(s+b_0)}(s+b_0)^n\, ds}{\int_0^\infty e^{-(s+b_0)}(s+b_0)^n\, ds}. \tag{13}$$

Making a change of variables of $x = s + b_0$ and using that

$$\int_{x_1}^{x_2} x^m e^{-x}\, dx = -e^{-x} \sum_{r=0}^{m} \frac{m! x^{(m-r)}}{(m-r)!} \Bigg|_{x=x_1}^{x_2} , \tag{14}$$

with $x_1 = b_0$ and $x_2 = s_{lim} + b_0$ for the numerator in (13) and $x_1 = b_0$ and $x_2 = \infty$ for the denominator, gives eq. (28.40) of Ref. [1],

$$\mathrm{CL} = 1 - \frac{e^{-(s_{lim}+b_0)} \sum_{r=0}^{n} \frac{(s_{lim}+b_0)^r}{r!}}{e^{-b_0} \sum_{r=0}^{n} \frac{b_0^r}{r!}} . \tag{15}$$

3 The Priors

There are two classes of priors in this method. One class contains Γ, the other the *nuisance variables*: ϵ, λ and b. In the former case, Γ is the parameter we wish to estimate with the experiment, in the latter, the parameters appear in (2) but are not directly interesting and are marginalized out of the final result. In both cases we have some minimum knowledge: We know with certainty some region where the parameter is physically allowed. So at a minimum, we can exclude non-physical regions by assigning them zero prior probability. In addition, for the nuisance parameters we have significant prior information. In fact, we have explicitly performed some other experiment in order to estimate their values. However, these estimations are imperfect, due to systematic and statistical uncertainties. The systematic uncertainty is assumed to result in a Gaussian distribution, while statistical uncertainty results in a distribution appropriate for the particular statistical process. As each prior is discussed in turn, note that the individual normalizations of the priors are ignored as they would cancel due to the application of (7).

3.1 The Decay Rate Prior

The decay rate is the parameter which this experiment attempts to limit. Because of this, we do not want to bias the result by using a prior with anything but the minimum of information. Such a prior is dubbed a *least informative prior* (LIP). The study of LIPs is ongoing especially in the case of continuous distributions. However, the currently accepted method of uniquely and objectively determining the LIP is by maximizing the *entropy* (MAXENT) of the parameter subject to any known constraints including any group symmetries appropriate to the problem [7]. In the case where it is known that a parameter must be strictly positive, nonzero the LIP determined by MAXENT is the so called Jeffreys's prior ($\mathbf{P}(x \,|\, I) \propto 1/x$). It is least informative because it gives no bias to any particular scale for x. In comparison, a uniform prior for x, which naively seems less informative, actually biases the scale towards higher values of x. Indeed with a uniform prior (eg. eq. (16)), $\mathbf{P}(x > x_0) > \mathbf{P}(x < x_0)$ for all x_0. However, proton decay lifetime limits have historically been calculated with the above mentioned simple Poisson limit [1], and this method, as seen in section 2.1, implicitly assumes a uniform prior. This same (non-LIP) uniform prior will be taken here in order that this limit would reduce to the historical one in the absence of uncertainties. Using a Jeffreys prior for the decay rate would result a lifetime limit that is numerically larger but, of course, with a different meaning.

Explicitly then, the prior probability distribution for the decay rate is,

$$\mathbf{P}(\Gamma\,|\,I) \propto \begin{cases} 1 & \Gamma \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

In practice, the calculation is truncated at very high values of Γ in order to make the integrals calculable.

3.2 The Efficiency Prior

The efficiency is estimated by generating many Monte Carlo events which simulate the decay of protons and the tracking of and the detector response to the decay products. The number of proton decay MC events passing the cuts divided by the number generated within the fiducial volume of the detector is then the efficiency. For $p \to e^+ \pi^0$, this estimate suffers uncertainty primarily due to our lack of knowledge of how the π^0 should be transported through the nucleus, as well as systematic differences between data and MC in the performance of vertex and direction fitters near the fiducial cut. This lack of knowledge has been estimated to be about 15% of the estimated efficiency, which is 45% to 60% depending on the exact analysis. So, prior to doing the experiment, we believe that the probability density of the true efficiency is Gaussian distributed about the estimated value. In addition, we are certain that the efficiency is above zero and is below one. This leads to a truncated Gaussian prior distribution for the efficiency,

$$\mathbf{P}(\epsilon\,|\,I) \propto \begin{cases} e^{-(\epsilon-\epsilon_0)^2/2\sigma_\epsilon} & 0 \leq \epsilon \leq 1 \\ 0 & \text{otherwise} \end{cases}, \tag{17}$$

where ϵ_0 is the estimated mean of the efficiency and σ_ϵ is the estimated uncertainty in that mean.

3.3 The Exposure Prior

The exposure is easily estimated given the detector content (ie. water), fiducial size, and live time of our data sample. The fiducial volume and live time are known to better than 0.1%. So while this prior is in principal a Gaussian, for this calculation it is taken as a Dirac δ-function,

$$\mathbf{P}(\lambda\,|\,I) \propto \delta(\lambda - \lambda_0), \tag{18}$$

centered at the estimated exposure, λ_0.

3.4 The background Prior

The true, real background is estimated by generating a sample of atmospheric neutrino Monte Carlo events, and counting the number which pass the proton decay selection criteria. At first look this seems very straight forward. However, this estimation itself is a counting experiment and the number of events counted will in general be a statistical fluctuation from the true MC background mean. Furthermore, the MC only simulates reality and because of uncertainties in the atmospheric neutrino flux and neutrino–oxygen cross sections used in the MC, the true MC background mean is systematically different from the true real background mean. Given this uncertainty we can ask what the probability that the true MC background mean is some value given the true real background mean. These concepts suggest the following derivation for the background prior which is similar to calculations used in other proton decay searches [6].

To calculate $\mathbf{P}(b\,|\,I)$, it is necessary to explicitly state some of the conditions represented by I. Namely the conditions that we count n_b atmospheric neutrino MC events passing the proton decay selection criteria out of C times more MC events than data events. Symbolically this can be written as, $I \equiv n_b C I'$, where I' symbolizes all conditions except those related to analyzing the MC sample. The method of section 2 is then followed so that the probability density function used as a background prior when analyzing the data is actually the posterior probability from the analysis of the MC "experiment".

So, by applying Bayes' theorem,

$$\mathbf{P}(b\,|\,I) = \mathbf{P}(b\,|\,n_b C I') \propto \mathbf{P}(n_b\,|\,b C I')\mathbf{P}(b\,|\,I'). \tag{19}$$

Here, $\mathbf{P}(b\,|\,I')$, is the information we have prior to considering the atmospheric neutrino MC. Since we are not constrained to match a historical background prior (the historical method assumes an exactly known background) the least informative Jeffrey's prior is chosen for $\mathbf{P}(b\,|\,I')$. This will give a numerically lower lifetime limit than if a uniform prior is chosen.

There is no direct way to calculate $\mathbf{P}(n_b\,|\,b C I')$, as the number of atmospheric neutrino MC events passing proton decay selections do not directly relate to the true, real mean background. However, it is related to the true mean atmospheric neutrino MC background, and this in turn is related to the true mean real background. In order to include these dependencies the partition theorem [1] is used. This is briefly described. Given an event, (possible outcome of an experiment), y, from a sample space, S, and the set of continuous values, $\{x\}$, representing unique events which partition S such that the union of all the members of $\{x\}$ is S, the partition theorem says,

$$\mathbf{P}(y\,|\,I) = \int \mathbf{P}(y\,|\,x I)\mathbf{P}(x\,|\,I)\; dx. \tag{20}$$

This is analogous to the completeness theorem (ie. the inserting of $\mathbf{1} = \int |x\rangle\langle x|\; dx$) of Quantum Mechanics.

In this case of estimating the background, x is chosen to be the parameter b_{MC}, the true mean MC background. And equation (19) becomes,

$$\mathbf{P}(b\,|\,I) \propto \mathbf{P}(b\,|\,I') \int_0^\infty \mathbf{P}(n_b\,|\,b_{MC} bCI')\mathbf{P}(b_{MC}\,|\,bCI')\, db_{MC}. \tag{21}$$

The first term, $\mathbf{P}(n_b\,|\,b_{MC} bCI')$, contains the statistical uncertainty in the background estimation due to finite size MC sample. Note that while b appears in this term due to the algebra, the probability of getting n_b events does not depend on b. This statistical uncertainty is expressed via the Poisson distribution. The second term, $\mathbf{P}(b_{MC} \mid bCI')$, contains our systematic uncertainty in the the MC background mean. This distribution is taken as Gaussian. The full background prior is then,

$$\mathbf{P}(b\,|\,I) \propto \frac{1}{b} \int_0^\infty \frac{e^{-b_{MC}} {b_{MC}}^{n_b}}{n_b!} e^{-\frac{(b_{MC}-bC)^2}{2\sigma_b^2}}\, db_{MC}, \tag{22}$$

for $b > 0$ and zero otherwise.

4 Example Applications of This Method

In the case of a proton decay analysis which has selection criteria such that the uncertainty in estimating the other parameters in (2) is small, we expect a limit which is close to the simple Poisson limit. Such a case is useful to check this method as well as it's implementation. This gives confidence when the method is used in analyses which can not be handled by simple methods.

The primary analysis for Super-Kamiokande's first $p \rightarrow e^+ \pi^0$ proton decay result has almost negligible background and small uncertainty in all parameters except, perhaps, the efficiency. It will be used as a low background example of applying this method.

The secondary analysis for Super-Kamiokande's first $p \rightarrow e^+ \pi^0$ proton decay result has very loose selection criteria. This increases the exposure as well as slightly reducing systematic errors associated with track fitting and energy scale. The dominant systematic in the efficiency is still due to uncertainties in the nuclear–pion model. However, these benefits are (over) balanced by the drawback of a much higher estimated background.

4.1 Low Background Limit

The first primary $p \rightarrow e^+ \pi^0$ result from Super–Kamiokande has an estimated efficiency,

$$\epsilon_0 = 0.44 \pm 0.06\,[0.0, 1.0]. \tag{23}$$

The values in the brackets give the region were the parameter is allowed to have non-zero probability. The estimated uncertainty is used as σ_ϵ in (17). The exposure is estimated to be,

$$\lambda_0 = 8.53 \pm \mathit{negligible}, \tag{24}$$

in units of 10^{33} proton · years. In the atmospheric neutrino MC sample, 3 events pass the selection criteria. With an oversampling of 30, the background is estimated to be,

$$b_{real} = 0.1 \pm 0.02\ [0.0, 1.0], \tag{25}$$

where the 20% uncertainty here is systematic. The rate prior is uniform and limited between 0 and 4 proton decays per 10^{33} years. As mentioned above, this limiting is to do with the pragmatics of the calculation. There is negligible probability for more than 4 proton decays per 10^{33} years given these data. The result is a 90% confidence limit for the partial lifetime of 1.56×10^{33} years. This can be compared to a simple Poisson background subtracted limit of 1.63×10^{33} years.

4.2 High Background Limit

The secondary result, used as consistency check with the first, has the following estimated parameters:

$$\epsilon_0 = 0.61 \pm 0.08\ [0.0, 1.0]. \tag{26}$$

$$\lambda_0 = 8.53 \pm negligable, \tag{27}$$

$$b_{real} = 3.5 \pm 0.7\ [0.0, 10.0], \tag{28}$$

As before, the exposure is in units of 10^{33} proton · years. The background is found to consist of 21 events passing the selection criteria, with an oversampling factor of 6.

Along with a higher background comes a non-zero number of candidate events, in this case 4, which is in keeping with the estimated background rate. This gives a partial lifetime limit of 0.97×10^{33} years. The simple Poisson background subtracted limit is 1.03×10^{33} years.

5 Conclusion

By employing Bayes' theorem and prior probability distributions it is possible to include our uncertainties into the calculation of a lower limit for the proton lifetime. The resulting limit is similar to what one would obtain with the simple Poisson limit and exactly reduces to this limit in the case of no systematic uncertainties. However, by including systematic uncertainties the lower limit on the lifetime will be lower than that found with out including systematics.

References

[1] R. M. Barnett *et al.*, Phys. Rev. **D54**, 1, (1996)

[2] T.J. Loredo, in *Maximum Entropy and Bayesian Methods*, edited by P.F. Fougère (Kluwer Academic, Dordrecht, The Netherlands, 1990). Available from: http://d0server1.fnal.gov/users/paterno/public_html/probability/

[3] G. D'Agostini, LANL preprint: hep-ph/9512295.

[4] M. Shiozawa, *et al.*, Search for Proton Decay via $p \to e^+ \pi^0$ in a Large Water Cherenkov Detector. Submitted to Phys. Rev. Lett.

[5] I. Bertram, *et al.*, A Recipe for the Construction of Confidence Limits, D0 Note 2775a. Available from: http://d0server1.fnal.gov/users/paterno/public_html/probability/

[6] C. McGrew, *et al.* A Search for Nucleon Decay using the IMB-3 Detector. To be published in Phys. Rev. D.

[7] E. T. Jaynes, in *E. T. Jaynes: Papers On Probability, Statistics, and Statistical Physics* , edited by R.D. Rosenkrantz (Kluwer, Boston, 1983)