Письма в ЭЧАЯ. 2015. Т. 12, № 2(193). С. 389–399

ФИЗИКА ЭЛЕМЕНТАРНЫХ ЧАСТИЦ И АТОМНОГО ЯДРА. ТЕОРИЯ

RADIOACTIVITY. CASE: RARE EVENTS

V. B. Zlokazov [1]

Joint Institute for Nuclear Research, Dubna

The paper discusses further development of the approach published in "Comp. Phys. Commun." (2014).

Low statistics means a little bit of information about the object of interest, so that a more or less exact parameter estimation and reliable statistical tests can be only a matter of chance, especially in the case of the exponential distribution, which is more intolerant to small samples (1–4 events) than the majority of other important distributions.

Therefore, the problem of optimization of the statistical analysis is especially actual for the exponentially distributed data, and the paper suggests, for both the parameter (mean) estimation and the statistical tests, a concept of a confidence interval, based on the order statistics, which, on the one hand, provides its clear and natural interpretation, and, on the other hand, is an optimum compromise between the criteria: "the shortest interval length" – "the largest size of the probability".

В работе обсуждается дальнейшее развитие подхода, опубликованного в «Comp. Phys. Commun.» (2014).

Малая статистика означает малое количество информации об объекте интереса, так что более или менее точное оценивание параметров и надежное статистическое тестирование могут быть лишь делом случая, особенно для экспоненциального распределения, которое более нетерпимо к малым выборкам (1–4 событий), чем большинство других важных распределений.

Поэтому проблема оптимизации статистического анализа особенно актуальна для экспоненциально распределенных данных, и в работе предлагается для оценки параметров и проверки гипотез понятие доверительного интервала, основанного на порядковых статистиках, который, с одной стороны, обеспечивает его ясную и естественную интерпретацию, а с другой стороны, оптимальный компромисс между критериями «кратчайшая длина интервала» – «максимальный размер вероятности».

PACS: 07.05.-t; 25.70.-z

INTRODUCTION

This paper is a further development of the approach published in "Comp. Phys. Commun." [1].

An exponential distribution (ED) plays a very conspicuous role in the experiments dealing with the radioactivity. Among them the most advanced ones, e.g., such as the synthesis of superheavy elements or the like ones, are characterized by a very small output, so that the information about the physical meaning of the observed process should be derived only from these scarce data.

[1] E-mail: zlokazov@jinr.ru

Generally, if the observed data contain a little bit of information, there are only three means to overcome this defect:

• large statistics of the data;

• superefficient estimation. It is the case when the accuracy of the unbiased estimate of the mean, based on m events, depends not on $1/m$ (as in usual efficient case), but on $1/m^2$. The former means: 4 times more events — 2 times better the accuracy. The latter: 2 times more events — 2 times better the accuracy — this is very profitable for the low statistics;

• a lucky chance — if the registered data are close (by accident) to the parameter of interest (usually the mean) of the distribution.

The first point is excluded from our study; the second one applies only to the uniform distribution. Thus, only the third one remains at our disposal. Let us call a distribution tolerant to the low statistics, if

1) it has a finite variance;

2) it has a property: any event falls into a Δ long vicinity of the mean with a greater probability than into any other interval of the Δ size (Δ is an arbitrary value).

1. THE MAIN DISTRIBUTIONS, WHICH TOLERATE THE LOW STATISTICS

Let the expectation of a random quantity be the parameter of interest. Then, the following distributions tolerate the low statistics.

• The normal distribution. Its probability density function is

$$p(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-c)^2}{2\sigma^2}\right). \tag{1}$$

Here the center c is the parameter of interest. For any time interval of a however small length δ, containing c, we see that the probability $\int\limits_{c-\delta}^{c+\delta} p(t)\,dt$ that our event falls into this interval is the greatest. It means that for experiments with low statistics the normal distribution is rather favorable — we have here the greatest chances that the events will be closely spread around the mean c, even if there are only few of them.

This gives us a possibility to define the low statistics formally. Referring to the widely spread semi-empiric opinion that in practice the average of 5 and more random values has already approximately the normal distribution, we can suggest that the data have a low statistics if it consists of not more than 4 items.

• The Poisson distribution. It is a distribution of a discrete random integer-valued variable ξ:

$$P(\xi = n) = \frac{a^n}{n!} \exp(-a), \tag{2}$$

where a (the parameter of interest) is both the mean and the variance.

The value n_x, where (2) is maximum, is close to a or, rather, to its nearest integer value. So, we see that (2) is also rather tolerant to the low statistics.

To a certain extent, the above definition of the tolerance to the low statistics is qualitative. One can invent densities, which formally satisfy it, but intuitively cannot be considered as tolerant, and, vice versa, one can invent such densities, which formally do not satisfy the above definition, but intuitively can be considered as tolerant. Examples are as follows:

1.

$$f(x) = \begin{cases} \sin^2(2\pi \cdot x/c) & \text{if} \quad x \in [c \cdot k, c \cdot (k+1)]; \\ \dfrac{\sin^2(2\pi \cdot x/c)}{(1+\epsilon)} & \text{otherwise,} \end{cases}$$

where $f(x)$ is defined in an interval of the x-axis of the length $c \cdot (2k+1)$, and ϵ is a small positive number.

2.

$$f(x) = \begin{cases} p & \text{if} \quad x \in [a, a+\epsilon]; \\ 0 & \text{otherwise,} \end{cases}$$

where $f(x)$ is defined in an interval of the x-axis $[0, L]$, a is an inner point of this interval, p is a constant, and ϵ is a small positive number, so that $a + \epsilon$ is much smaller than L.

However, the above definition conveys the idea of the tolerance to the low statistics, and gives reliable examples of tolerant distributions (the Poisson and Gauss ones), so that if a distribution is close to either of them in the sense of the C-metric, it can be counted tolerant.

2. THE EXPONENTIAL DISTRIBUTION

Unfortunately, the absolute majority of other widely-used distributions do not favor the low statistics, and among them the most striking example of the contrast between "the most probable" and "the most expected" is given by the exponential probability distribution.

The exponential distribution (ED(T)) for the quantity ξ with the parameter T is defined as follows:

$$F_\xi(t, T) = \begin{cases} 1 - \exp\left(\dfrac{-t}{T}\right) & \text{if} \quad t \geqslant 0; \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

Here t is the time. In the applications such a form of the T parameter is preferable, since in this case T (the decay constant) and t are measured in the same direct time units.

We have here the distribution density $p(t) = \exp(-t/T)/T$, which is nonzero valued in $[0, \infty)$ and T as the mean and T^2 as the variance.

At $t = 0$, the density $p(t)$ has the maximum and it means that the decays, however close to $t = 0$, are the most probable ones. In [1], it has been shown that while observing a radioactive decay, we have almost thrice more chances to observe a value close to 0 than to T.

It does not play an essential role if the statistics is large, but it may be of crucial importance if we have only few events.

A radioactive process looks like this — an avalanche of events at the beginning, and then the succession of a diminishing geometric progression of the rest. This is a contrast to the normal distribution.

3. THE GAMMA DISTRIBUTION

For an exponential random quantity ξ there is a distribution, which is closely connected with it. It is the one with the following density function:

$$g(t, m, T) = \begin{cases} \dfrac{t^{m-1}}{T^m (m-1)!} \exp\left(\dfrac{-t}{T}\right) & \text{for} \quad t \geqslant 0; \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where m is positive integer, and T is positive real. The mean of the distribution (4) is mT and the variance is mT^2.

For $m = 1$ the function (4) is the usual exponential probability distribution.

Let a sample of random values $t_1, t_2, \ldots, t_m$ of ξ be given, and consider the following quantities:

$$S = \sum_{i=1}^{m} t_i, \quad S_m = \frac{S}{m}. \tag{5}$$

The random quantity S has the (4) distribution (see, e.g., [2]). The density of the S_m distribution is $m \cdot g(mt, m, T)$, and its mean and the variance are equal to T and to T^2/m, respectively. The maximum of the density (let it be t_x) is reached at the root of the equation

$$m\left[\frac{(m-1)(mt)^{m-2}}{T^m(m-1)!} - \frac{(mt)^{m-1}}{T^{m+1}(m-1)!}\right]\exp\left(-\frac{mt}{T}\right) = 0,$$

from which we obtain $t_x = (m-1)T/m$.

For the case of low statistics $(m = 1, 2, 3, 4)$ we see that this maximum is rather far from the mean T. For instance, if $m = 2$, the distances between 0 and t_x, and between t_x and T, are equal to $T/2$, i.e., for $m = 2$ the half-sum $(t_1 + t_2)/2$ has equal chances to be close to 0 as well as to T.

If $m = 3$, then the $(t_1 + t_2 + t_3)/3$ has two chances against one that it will be closer to T than to 0; and so on: $m - 1$ chances for "T" against one chance for "zero". While $m \to \infty$, according to the Central Limit Theorem, the distribution (4) tends to the normal one with the center T.

Summarizing, we can say that the gamma distribution is not tolerant to the extremely low statistics $(m = 1, 2, 3, 4)$.

4. THE PROBLEMS

Given a random sample $S = t_i,\ i = 1, 2, \ldots, m$ of size m from an ED (the times of a radioactive decay), we can specify the following tasks of their analysis:

1) on the basis of S estimate the T parameter and its accuracy;

2) for the given T test the hypotheses:

a) does each of $t_i,\ t_i \in S$ correspond to the model $F(t, T)$?

b) has the whole set S the distribution $F(t, T)$?

We shall start with the second problem, because for the rare events one can get more reliable results for the statistical tests rather than for the parameter estimates.

To make a decision on the correspondence of the set S to $F(t, T)$, it is necessary to build a CI — a confidence interval (in the decision-making called also critical region); it is an interval $[a, b]$ on the t-axis, into which the tested values of our random variable t_i (case (a)) or some function s of the set S (statistic) (case (b)) fall with a certain confidence probability (P_c); if the event t_i or the statistic s fall into $[a, b]$, then they do not contradict the tested hypothesis that the distribution is really $F(t, T)$ (but, of course, do not yet confirm it).

As a rule, use is made of a two-sided CI $[M \pm \sigma]$, where M is the mean value and σ is the square root of the variance.

For the Gaussian distribution this corresponds to $P_c \approx 0.68$, and for such a test the ratio of the "pro" and "contra" chances is equal to approximately two.

However, in our case, one-sided CIs are also of great interest [1], when, e.g., $m = 1$, i.e., for the problem 2 (a). These CIs have the form $[0, 2T]$, where T is the tested value of the ED parameter, since in case of the ED events, which are close to 0, occur with the maximum probability, and, of course, 0 should be the lowest bound of such a CI.

Remark. The lowest CI bound in case of hypothesis testing should not be confused with the lowest CI bound in parameter estimation. In the latter case, a CI $[T_{\min}, T_{\max}]$ describes with a certain confidence probability the most probable values of the T parameter, and, of course, $T_{\min}$ is always greater than 0.

In case of hypothesis testing, a CI $[t_{\min}, t_{\max}]$ describes with a certain confidence probability the most probable t values for the tested T parameter, and, therefore, $t_{\min}$ can be equal to 0.

A two-sided CI for the testing hypotheses is appropriate, if 0 is not the value of the maximum probability density.

5. OPTIMIZATION OF THE CONFIDENCE INTERVAL

For a given $F(t, T)$ we shall use a concept of an optimal confidence interval $[a, b]$ (OCI) described in [1]. Such an OCI should have minimal difference $b - a$, and, at the same time, the probability of the events to belong to the interval $[a, b]$ "pro chances" should be maximum; since these conditions contradict each other, an OCI is one of the two compromises:

— for a given length $b - a$ find an interval with the best ratio "pro/contra";

— for a given ratio "pro/contra" find an interval of the shortest length $b - a$.

Apart from this, the physical meaning of the interval $[a, b]$ and its bounds a and b should be clear and natural.

For an exponential distribution $F(t, T)$ and $m = 1$ one can propose a semi-empiric approach, which would allow us to build such a one-sided OCI (i.e., $[0, 2T]$) with a minimum of arbitrary assumptions about the data [1].

Let us see what can be done for the case of two-sided CIs ($m > 1$). Let σ be the square root of the S_m variance. Then, the usual two-sided CI is $[T - \sigma, T + \sigma]$. It is a fixed compromise between the size of the CI and the area of the total probability covering it.

However, it is not clear how this probability is distributed within the CI — generally this CI does not reflect the structure of the ED, in particular, its asymmetry. Thus, its physical meaning is often not clear. Therefore, it would seem desirable to elaborate a scheme of a CI, which would keep the advantages of the usual CI and be free of its drawbacks.

6. ORDER STATISTICS

For this reason, let us make use of the so-called order statistics. The method based on them is, in our case of an ED, especially convenient, because they can be represented as easily integrable analytical functions.

Let the items t_i of a sample S be arranged in an increasing order. Following [2], we define the following order statistics:

1) denote the minimal value in the sample S as u_1; it is a random quantity with the probability density

$$g_1(u_1) = \frac{m}{T} \exp\left(-m \cdot \frac{u_1}{T}\right) \quad \text{for} \quad u_1 \geqslant 0;$$

2) and denote the maximum value in the sample S as u_m; it is a random quantity with the probability density

$$g_m(u_m) = \frac{m}{T} \exp\left(-\frac{u_m}{T}\right)\left(1 - \exp\left(-\frac{u_m}{T}\right)\right)^{m-1} \quad \text{for} \quad u_m \geqslant 0.$$

Omitting integrations, which can be easily reconstructed, we get the expectations $\hat{E}u_1$ and $\hat{E}u_m$ for the cases of low statistics, i.e., for $m = 2, 3, 4$ (Table 1).

To compare a UCI — usual confidence interval $[T - \sqrt{m}T, T + \sqrt{m}T]$ and an OCI $[T_{\min}, T_{\max}]$, let us consider the following two tables (Tables 2 and 3) for the different T; one can see that the results weakly depend on the parameter T (certainly, excepting the interval length).

Here Prob "pro" is the probability to accept the hypothesis, if the tested value falls into the CI.

The analysis of these tables allows us to make such conclusions.

• The OCIs really have a special psychological advantage — they have the most clear interpretation as bounds between the most typical minimal and the most typical maximum values of the random quantity.

• For $m = 2$ the probability covering the OCI may seem to be too small; in this case, it is more appropriate to solve the following optimization problem: for the fixed probability (e.g., 0.68) find the shortest CI.

Table 1. **Expectations of the order statistics**

Parameter	$m = 2$	$m = 3$	$m = 4$
$T_{\min}$	$T/2$	$T/3$	$T/4$
$T_{\max}$	$(3/2)T$	$(11/6)T$	$(25/12)T$
Length	T	$(3/2)T$	$(11/6)T$

Table 2. $T = 20$

Names	OCI ($m = 2$)	OCI3	OCI4	UCI ($m = 2$)	UCI3	UCI4
Prob "pro"	0.55	0.83	0.95	0.74	0.75	0.71
Ratio "pro/contra"	1.2	5.0	19.9	2.9	2.9	2.5
CI length	T	$1.5T$	$1.835T$	$1.41T$	$1.41T$	$1.41T$

Table 3. $T = 80$

Names	OCI ($m = 2$)	OCI3	OCI4	UCI ($m = 2$)	UCI3	UCI4
Prob "pro"	0.54	0.83	0.95	0.75	0.72	0.71
Ratio "pro/contra"	1.2	5.0	18.4	2.9	2.5	2.4
CI length	T	$1.5T$	$1.83T$	$1.41T$	$1.41T$	$1.41T$

• For $m = 3, 4$ the optimization is: among the intervals with the lengths $(3/2)T$ and $(11/6)T$, respectively, find those having the greatest covering probability.

In all the cases, to keep the clearness, they should have either a as the 1st order statistics or b as the maximum order one, or both should be those of OCIs (see the example in Sec. 10).

7. PARAMETER ESTIMATION

In the case of an ED and data with low statistics, this problem requires a special consideration. The usually used maximum likelihood estimator (MLE) is the average $\hat{T} = S_m$ given by (5), which, for the case of one event, is the data t_1 itself.

• Case of one event. The MLE is based on an assumption that on the average the data likelihood is maximum, that here turns out to be false. In the case of one event t_1, it is more reasonable to consider t_1 as an estimate of the lower bound for T, the argumentation being as follows.

The probability P_k of an inequality $t_1 < kT$ is equal to $P_k = 1 - \exp(-k)$; here k is an arbitrary number. We can try different estimates of T, which still guarantee that the inequality holds. The minimal of them is obviously $\hat{T} = t_1/k$. It is the estimate of the lower bound of T with the confidence probability P_k, which depends on k. For $k = 2$ $P_2 \cong 0.865$; For $k = 1$ $P_1 \cong 0.63$.

• Case of m events, $m = 2, 3, 4$. The estimate of T is S_m (the average), and it is appropriate to take as bounds the same OCI, based on order statistics $[T_{\min}, T_{\max}]$, for the same reasons as in case of the hypothesis testing. It provides a better compromise between the CI length and the probability covering it than the UCI does.

8. HYPOTHESIS DISCRIMINATION

Hypotheses testing gives us an answer to a question: Can the tested value originate from the tested distribution? But it gives no answer to the question: Does the tested value originate from the tested distribution?

Such answers can be obtained using the techniques of the hypothesis discrimination. In our case, we can proceed in the following way.

In principle, the problem can be solved by testing a finite number of hypotheses exhausting all the realistic interpretations of our data (if it is possible) and selecting only one, which does not contradict the data, while all the other do. Certainly, in our case of low statistics, a more or less reliable discrimination can be made of not more than two hypotheses.

So, we have the two hypotheses — H_0 : that $T = T_0$, and H_A : that $T = T_A$, let $T_A > T_0$.

• Case of one t_i. We shall use the OCIs for one event described in [1]. Events from the interval $[0, 2T_0]$ do not contradict the both hypotheses; events from $[2T_0, 2T_A]$ contradict only H_0, and events from $[2T_A, \infty]$ contradict the both H_0 and H_A. The critical region is the interval $C_r = [2T_0, 2T_A]$. If $t_1 \in C_r$, we accept H_A and reject H_0. The Type I error (to reject the true hypothesis) is equal to $\int\limits_{2T_0}^{2T_A} \exp(-t/T_0)/T_0 \, dt$, if H_0 is true, and the Type II error (to accept the false one) is the same.

• Case of several t_i. The average S_m (5) has the $m \cdot g(mt, m, T)$ distribution, where $g(t, m, T)$ is the gamma distribution (4). Taking the order statistics intervals $[T_{0\,\min}, T_{0\,\max}]$ and $[T_{A\,\min}, T_{A\,\max}]$ as OCIs, we can build the critical region for the discrimination of H_0 and H_A. For the simplicity reason suppose that $T_{0\,\max} > T_{A\,\min}$.

Let us use the following notation:

$$a_1 = T_{0\,\min}, \quad a_2 = T_{0\,\max}, \quad b_1 = T_{A\,\min}, \quad b_2 = T_{A\,\max}.$$

We can divide the whole t-axis into the following intervals:

$$R_1 = [0, a_1], \quad R_2 = [a_1, b_1], \quad R_3 = [b_1, a_2], \quad R_4 = [a_2, b_2], \quad R_5 = [b_2, \infty],$$

and set up the following rules for the decision-making:

1) if S_m falls into R_1 or R_5, the data contradict the both hypotheses;

2) if S_m falls into R_2, H_0 is accepted, and H_A is rejected;

3) if S_m falls into R_4, H_0 is rejected, and H_A is accepted;

4) if S_m falls into R_3, the hypotheses cannot be distinguished (for this statistical level), both the hypotheses can be accepted.

The Figure can illustrate this case.

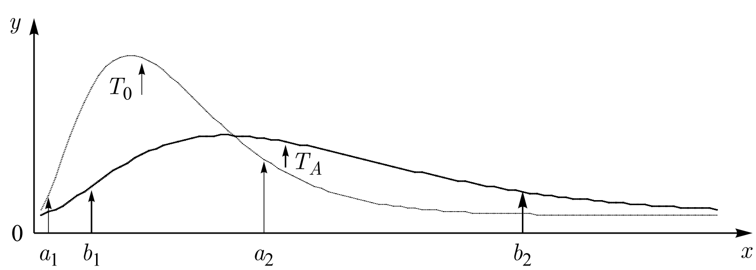
The Type I error is $\int\limits_{T_{A\,\min}, T_{A\,\max}} \exp(-t/T_0)/T_0\, dt$ and the Type II error is $\int\limits_{T_{0\,\min}}^{T_{0\,\max}} \exp(-t/T_A)/T_A\, dt$.

From Table 1 we can get the values of the order statistics for $m = 3$,

$$T_{0\,\min} = \frac{T_0}{3}, \quad T_{0\,\max} = \frac{11}{6} T_0, \quad T_{A\,\min} = \frac{T_A}{3}, \quad T_{A\,\max} = \frac{11}{6} T_A.$$

Calculating the integrals of the Type I and II errors, we shall get: $T_{\rm I}$ error ≈ 0.488 and $T_{\rm II}$ error ≈ 0.447. Chances to discriminate the hypotheses for $m = 3$ and the ratio $T_A/T_0 = 2$, given by these probabilities, are not too large. However, if $T_A/T_0 = 3$, the corresponding probabilities are 0.364, 0.352 and the chances increase, even if m remains the same.

We can build a function $f(R = T_A/T_0)$, which describes the dependence of Type I and II errors on R and estimate the optimum R, for which the hypotheses can be discriminated with acceptable error probabilities.



The gamma distribution $m = 3$. The confidence intervals $[a_1, a_2]$ (thin line) and $[b_1, b_2]$ (thick line) for the discrimination of the hypotheses $T_0 = 20$ and $T_A = 40$

9. RADIOACTIVITY AS TIME PROCESS

There are two types of registration of a radioactive decays proceeding in the time:

— that beginning at a definite point $t = 0$;

— that performed within a finite time interval $[t_1, t_2]$, $t_1 \neq 0$ and $t_2 \neq \infty$.

The first model suggests that at a certain moment a decaying mass emerges at once; the second one is more complicated for the analysis and we consider it here.

As mentioned above, a radioactive process is an avalanche of events at the starting moment of the measurement and then the succession of a diminishing geometric progression of the resting decays. Suppose that the decaying mass has sufficiently many objects and is in an equilibrium state (no new objects appear). Theoretically, any two finite observation intervals, in which the number of registered events is greater than zero, contain the information about the decay constant T. In [1], it has been shown that if the decaying mass is sufficiently large, then for a however big decay constant the probability that at a finite point t_1 a decay will take place, tends to 1, if $t \to \infty$. But the accuracy and reliability of the T estimate are negatively influenced not only by the low statistics, but also by the trimming of the observation interval: $[t_1, t_2]$ instead of $[0, \infty]$.

Let us consider the second case in more detail. If the events are observed only within an interval $[t_1, t_2]$ (a limited subinterval of the whole t-axis) and the events falling in $[0, t_1]$ or $[t_2, \infty]$ are not registered, the distribution function is

$$F(t, T) = \begin{cases} 1 - \exp\left(-\dfrac{t}{T}\right) & \text{if} \quad t \in [t_1, t_2]; \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

We have the distribution density which is nonzero only in $[t_1, t_2]$, where it has the form

$$p(t) = \frac{\exp\left(-\dfrac{t}{T}\right)\Big/ T}{\exp\left(-\dfrac{t_1}{T}\right) - \exp\left(-\dfrac{t_2}{T}\right)}. \tag{7}$$

Its mean T_t is given by the formula

$$T_t = \frac{T\left(\exp\left(-\dfrac{t_1}{T}\right)\left(\dfrac{t_1}{T} + 1\right) - \exp\left(-\dfrac{t_2}{T}\right)\left(\dfrac{t_2}{T} + 1\right)\right)}{\exp\left(-\dfrac{t_1}{T}\right) - \exp\left(-\dfrac{t_2}{T}\right)}. \tag{8}$$

The mean (8) can differ very strongly from the “true” mean T (that is from the mean value corresponding to (3)), especially if the sample is trimmed in the vicinity of zero, and obtaining an accurate estimate of T is not an easy problem.

In case of low statistics, we cannot use histogram methods for the evaluation of T; the maximum likelihood estimator fails here too. Indeed, the likelihood function is

$$L(T) = \prod_{j=1}^{m} p(t_j), \tag{9}$$

and the maximum likelihood estimate of T is its value, at which (9) has the maximum. Substituting (7) in (9) for $p(t_j)$, we see that the maximum of (9) will be reached at $T = \infty$, if, at least, one t_j gets in the interval $[t_1, t_2]$.

Bearing in mind that among all functions, which could be taken as estimator, the best is the mean, let us study under what conditions the mean T_t of a sample from $[t_1, t_2]$ will be close to T?

We can use the fact that (8) depends on t_1, t_2, so that selecting the optimum t_1, t_2 (or only t_2, because t_1 is normally fixed), we have a chance to make T_t be equal to T.

Writing the equation

$$\frac{T\left(\exp\left(-\frac{t_1}{T}\right)\left(\frac{t_1}{T}+1\right) - \exp\left(-\frac{t_2}{T}\right)\left(\frac{t_2}{T}+1\right)\right)}{\exp\left(-\frac{t_1}{T}\right) - \exp\left(-\frac{t_2}{T}\right)} = T$$

and dividing the both parts by T and reducing the fraction by $\exp(-t_1/T)$, we shall get

$$\left(\frac{t_1}{T}+1\right) - \exp\left(-\frac{t_2-t_1}{T}\right)\left(\frac{t_2}{T}+1\right) = 1 - \exp\left(-\frac{t_2-t_1}{T}\right),$$

from which the needed condition follows as:

$$\ln\frac{t_2}{t_1} = \frac{t_2-t_1}{T}. \quad (10)$$

If t_1 is fixed and t_2 satisfies (10), the mean of the sample in $[t_1, t_2]$ is an unbiased consistent estimator of the T parameter.

Equation (8) can be also used as moment estimator of T, i.e., substituting the sample mean for T_t in (8) and solving it with respect to T, we shall get an estimate of T. The accuracy of this estimate can be evaluated by expanding (8) into linear terms of T, and deriving T from it as a function of T_t. However, the success of this method requires a large statistics.

10. TESTING

We can use the above-described method to the data, published in [3] in order to see to what extent the confidence intervals reported there are optimum and compare them with the results given by our method.

In the TASCA case, the authors reported the following estimate of the $^{294}117$ half-life (in ms): 51^{+94}_{-20}; the analysis used two decay chains; in the DGFRS case, the estimates, based on three chains, are as follows: 50^{+60}_{-18}.

It is very strange that these CIs are strongly asymmetric on the right from the $T_{1/2}$ point. The exponential distribution (as mentioned at the paper beginning) is an avalanche of the events in the first time period ($[0, T_{1/2}]$) — one half of the total decay integral — then one fourth in the next $[T_{1/2}, 2T_{1/2}]$, and so on. Therefore, the events to the right from $T_{1/2}$ are those of a small and rapidly diminishing probability, and, certainly, first of all, a CI should cover events from the left side.

The order statistics for $m = 2$ is $[25.5, 76.5]$ ms. The minimal CI for the probability 0.68 is $[8, 65]$. Comparing the CI lengths $94 + 20 = 114$ ms and $65 - 8 = 57$ ms, we see that the TASCA CI is not optimum — the same covering probability, but a much longer length (almost twice).

The order statistics for $m = 3$ (DGFRS case) is $[17, 92]$ ms. The covering probability is about 0.83. The minimal CI with the length 108 is $[12, 120]$, which is covered by the probability about 0.86. Here we see also that the DGFRS CI is not optimum — not only its CI length is longer, but also the covering probability is significantly smaller compared with that of the OCI.

CONCLUSION

It has been shown that, unlike the normal and Poisson distributions, the exponential one is very intolerant to the low statistics (1–4 events), so that the more or less exact parameter estimation and reliable statistical tests strictly require the optimized techniques.

As such, for both the parameter (mean) estimation and the statistical tests a concept of a confidence interval is formulated based on the order statistics, which, on the one hand, provides their clear and natural interpretation, and, on the other hand, means a good compromise between the criteria: "the shortest interval length" – "the largest size of the covering probability".

REFERENCES

1. *Zlokazov V. B.* Confidence Interval Optimization for Testing Hypotheses under Data with Low Statistics // Comp. Phys. Commun. 2014. V. 185. P. 933–938.
2. *Freund J. E.* Mathematical Statistics. Englewood Cliffs: Prentice-Hall, Inc, 1962. No. 3.
3. *Khuyagbaatar J. et al. (Collab.).* ^{48}Ca + ^{249}Bk Fusion Reaction Leading to Element $Z = 117$; Long-Lived α-Decaying ^{270}Db and Discovery of ^{266}Lr // Phys. Rev. Lett. 2014. V. 112. P. 172501.

Received on July 2, 2014.