RADIACTIVITY REGISTERED WITH A SMALL NUMBER OF EVENTS

Zlokazov V.B.[1] and Utyonkov V.K.[2]
JINR, LIT LYAR
[1] - zlokazov@jinr.ru [2] - utyonkov@jinr.ru

Due to lack of time we shall mention here only new results which are not contained in earlier papers.

The experiments on the synthesis of the superheavy elements yield chains of the sequential decay of the elements E_i, following a scheme: $E_i \to E_{i+1} \to E_{i+2} \to E_{i+3}$ etc. Each chain consists of registered decay times t_i of the corresponding element and some chains can be not complete.

These experiments can be carried out by different groups of scientists using different means of the chain production and registration and the goal of the analysis of produced data is

- estimate the statistical characteristics of the appearing isotopes and the accuracy of these estimates.
- test the compatibility of results of different experimenter groups.

We consider the case: 3 decay stages, 2 different groups.

At first we have to answer the question: should we analyze separate items in the chains or consider total chains as single items? To make it clear let us recall the axioms of the mathematics of radioactivity [1]

- any k decays belong to the same type, are mutually independent, occur in non-intersecting time points and their probability within any time interval $[t_0, t_0{+}\Delta t]$ does not depend on the choice of t_0;

- the decays are so called 'rare events' - the probability of the registration of 2 and more events for any small Δt is infinitesimal compared with the probability of registering 0 or 1;

- Markov property of decays - the probability of the registration of any number of decays after any moment t does not depend on the prehistory.

Thus our problem formally is: experimental events t_{iEc} being given where i = number of event, E = name of the isotope, c = number of the team conducting the experiment,

- estimate the decay constants of each group of data $t_{iEc}, i = 1, ..n_{Ec}$ for each couple Ec
- test these estimates $\hat{t}_{Ec}$ for each E and c for compatibility.

Some event registration methods can break some of the above rules. For instance, if the third axiom is broken, we should analize not the separate events, but their chains (for each E and c), which complicates the analysis.

Therefore, at first we verify on data which today are considered as safely established, whether their events really follow the rule of uncorrelatedness of the events t_{iEc} of E_j and E_{j+1}.

Let 2 simulated samples of uncorrelated data $A_1 = t_{11}, t_{12}, .., t_{1n}$ and $A_2 = t_{21}, t_{22}, .., t_{2n}$ with the same exponential distribution function be given and let us calculate their sample correlation coefficient r

$$r = \sum_{i=1}^{n} (t_{1i} - m_1)(t_{2i} - m_2)/(\sigma_1 \cdot \sigma_2)/n \quad (1)$$

where $m_1, m_2, \sigma_1, \sigma_2$ are expectations and sigmas of the both distributions.

Theoretically in an ideal case (infinite n, absolutely exact $m_1, m_2, \sigma_1, \sigma_2$) we should have $r = 0$. But in our case of finite data the calculations of r won't give zero.

The Monte-Carlo tests showed that the main factor defining the range of sample values of r is the data size n. The Table 1 shows the mean and the minimal and maximum values of r in dependence on the statistics n.

	Min	Max	Mean	MAbs	2par	3par
n= 10	-0.1335	0.3646	0.0865	0.1132	0.0942	0.1291
n= 20	-0.3781	0.2673	-0.0200	0.1024	0.1636	0.0779
n= 30	-0.1011	0.1312	0.0054	0.0367	0.0709	0.0554
n= 40	-0.2348	0.3436	0.0546	0.1026	0.0530	0.0539
n= 50	-0.1043	0.2531	0.0091	0.0489	0.0610	0.0639
n= 60	-0.1054	0.2027	0.0067	0.0406	0.0615	0.0299
n= 70	-0.0456	0.2101	0.0333	0.0456	0.0424	0.0613
n= 80	-0.0661	0.1659	0.0184	0.0375	0.0490	0.0315
n= 90	-0.2169	0.0826	-0.0043	0.0404	0.0539	0.0267
n= 100	-0.0623	0.1895	0.0269	0.0394	0.0250	0.0302

The data in the 2st through the 5th columns contain quantities obtained for the samples with the equal parameters, the 6th and the 7th columns - for the samples with parameters, multiplied by 2 and 3, respectively.

The analysis of this table shows that the estimates of correlation coefficients for the finite samples differ from zero (especially for $n = 10$ and 20).

They can be called the by-product effects; still, they can be used as measures of the true correlations, because they vary in a rather stable range and practically don't depend on parameters.

A rigorous mathematical analysis of the distribution of sample correlation coefficients can be found in [4, Exercises, p.591]., but it covers only assymptotics of the normal event distribution and is described by a very cumbersome formula (including infinite series), which can be hardly used in the practice.

However, based on the Table 1, we can point out an interval (e.g. $[-0.4, 0.4]$) so that if the obtained sample correlation coefficient does not exceed this range, the tested samples don't contradict the hypothesis about their uncorrelatedness.

Then to verify that the successive radioactive decays really are uncorrelated, we can use the data with a rather good statistics, e.g. $^{288}Mc \rightarrow \ ^{284}Nh \rightarrow \ ^{280}Rg$, 63 events for each decay step [3].

The performed analysis gave the estimate $r = -0.18$. So the real practice shows examples not rejecting the Markov property of the successive decays.

As for our data (11 events in each sample) in $^{115}Z \to {}^{113}Z$, we obtained $r = -0.11$ and for the samples $^{113}Z \to {}^{111}Z$, $r = +0.30$. So we can accept the hypothesis that also here the successive decays are uncorrelated and we can use for the comparison not the fixed chains of events, but separate events in each data group.

Next, we can estimate the parameters Ts of all the exponentials - the probability distributions of decay times t_i:

$$P(t_i \ < \ t) = P(t, T) = 1 - exp(-t/T) \quad (2)$$

where the expectation of t_i is T and its variance T^2.

The sums of times of m decays

$$S = \sum_{i=1}^{m} t_i. \quad S_m = S/m. \tag{3}$$

have the real (m, T)- gamma distribution [2]. It is the one with the following density function

$$g(t,m,T) = \begin{cases} \frac{t^{m-1}}{T^m(m-1)!} exp(-t/T) & \text{for } t \geq 0; \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

m = positive integer, T = positive real.
The expectation of (4) is mT and the variance mT^2. While m tends to the infinity, $g(t,m,T)$ tends to the normal density.

The statistic S (3) is a sufficient one for (2) [4] and S_m can serve as estimate of the parameter T. At the same time it is the maximum likelihood estimate (MLE) and is an efficient (having the variance assimptotics $1/n$) estimate of T.

The density of S_m is $m \cdot g(mt, m, T)$, and its mean and the variance are equal to T and to T^2/m, respectively.

From the maximum of the density (4) (let it be t_x) we obtain $t_x = (m-1)T/m$.
For the case of small m we see that this maximum can be rather far from the mean T.

Therefore, we can say that the gamma-distribution is not favourable to the low statistics - it is the case where 'the most probable is not the most expected'.

Now we return to the initial problem: The two sets of decays $A_1 = (t_{11}, t_{12}, .., t_{1m})$, $A_2 = (t_{21}, t_{22}, .., t_{2m})$ being given test the two hypotheses

1. $H_1 : A_1$ and A_2 have the same distribution;
2. $H_2 : A_1$ and A_2 have the different distributions.

In [7] a conventional approach to solve this problem is described: build the confidence intervals for the estimates of the means (the Student's tests) and of the variances the χ^2 test:

$$-t_{\alpha/2,m-1} < (\hat{T} - T)/\sigma) < t_{\alpha/2,m-1}$$

$$\chi^2_{1-\alpha/2,m-1} < ((m-1)\hat{s}/T^2) < \chi^2_{\alpha/2,m-1}$$

where $t_{\alpha,m-1}$ and χ^2.. are the values of the Student's and χ^2 standard random quantity, corresponding to the α probability and the number of degrees of freedom $m-1$, It is reasonable enough to accept H_1 if the intervals I_1 and I_2 overlap and the probability covering the interval of overlapping is sufficiently large or, calculating the so called Type I and Type II errors (reject a true hypothesis and accept the wrong one) and analyzing them, make the corresponding decision.

For a given $F(t,T)$ we shall now consider a concept of an optimal confidence interval $[a,b]$ (OCI) for testing the hypotheses. Such an OCI should have the maximum accuracy of the testing criterion which means: the minimal difference $b-a$ and at the same time the probability of the events to belong to the interval $[a,b]$ should be maximum.

Since these conditions contradict each other an OCI is one of the two compromises: [5]

- 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 $b - a$.

Very often for the normal distribution one takes the interval $[T-\sigma, T+\sigma]$, which for the gamma distribution means $[T - T/\sqrt{n}, T + T/\sqrt{n}]$ and is a rather good compromise between the estimate accuracy and probability cover; but the convergence of the gamma distribution to the normal one is very slow.

Apart from these criteria we can build an OCI from a requirement that the physical meaning of the interval $[a, b]$ and its bounds a and b should be maximum clear and natural.

One of the ways to build such an OCI is making use of the so called order statistics, which for the distribution (4) can be derived as easily integrable analytical functions.

Following [2] we define the following order statistics. Denote the minimal value in the sample S as u_1; and denote the maximum value in the sample S as u_m; their expectations $\hat{u}_1 = \hat{E}u_1$ and $\hat{u}_m = \hat{E}u_m$ which are the order statistics of the 1st and the last rang, the meaning of which is

1. $\hat{u}_1$ is the average minimal value in a sample;
2. $\hat{u}_m$ is the average maximum value in a sample.

They can be used as the most 'natural' bounds of the average decay time in the sample.

An additional means to test the 2 distributions for congruence is to take the expectation E_d and the variance (sigma) σ_d of the normalized differences of the both distributions t_{1i} and t_{2i} and build the confidence interval $[\sigma, \sigma]$ and see whether the E_d falls in it.

Unfortunately, the formula for the variance of the difference of 2 gamma-distributed random quantities has no simple closed form, so that it is difficult to estimate the trustworthness of this test.

And now we describe a method to solve a very important problem: is the analyzed data really a sample from a single exponential distribution or from a mixture?

This method is based on checking the sample relations between such fundamental characteristics of any event distribution as the mode, the median and the mean. In case of exponentials these relations are uniquely defined and are different for single components and mixtures.

The relation between the mean E and the median M for a single exponential is

$$M = E \cdot ln(2).$$

For the ratio K of 'median / mean' in finite samples (e.g. n=15) the statistical tests gave

$$K = 0.7 \pm 0.17$$

where the confidence interval has the size of 67%.

However, if the data is the decay of two (or more) sources, depending on the structure of the mixture, K can differ from $ln(2)$ by several orders. Therefore, K is both a simple and a rather reliable indicator of whether the source is pure, or it has an asymmetric admixture.

To conclude the report let us build the exponential distribution independent on the parameter T. The sufficiency of (3) means that we can transform the distribution (2) to a form which is not explicitely dependent on the parameter T.

Let's formulate a Theorem:
A set of decays $A = t_i, \ i = 1, 2, .., m$ being given, the values $m \cdot t_i/\hat{\tau}$, where $\hat{\tau} = \Sigma_{i=1}^{m} t_i$, will have asymptotically the exponential distribution $1 - exp(-t)$.

The proof. Let's denote $y = m \cdot t / \Sigma_{i=1}^{m} t_i$. Then for arbitrary x $P(y < x) = P(t < xm \cdot \Sigma_{i=1}^{m} t_i) = 1 - exp(-x \cdot \Sigma_{i=1}^{m} t_i/m/T) \to 1 - exp(-x)$ while $m \to \infty$. The latter has the unit as parameter, and as seen, is asymptotically T-independent along with its expectation and variance, which are equal to the unit.

This enables us to deal not with the gamma-distribution, but with the simpler exponential one and get the scale-independent quantities by analytical operations. To get such quantities for a concrete event set A one can simply multiply the independent ones with $\hat{\tau}$.

And here let us see relations 'CI length' - 'covering probability', using the parameter independent distribution $1 - exp(-t)$.

m	u1	um	pr1	-s	+s	pr2	os1	+s	pr2
1	0.500	1.500	0.5391	0.000	2.000	0.9103	0.500	2.000	0.6392
2	0.333	1.833	0.8307	0.293	1.707	0.8272	0.333	1.707	0.8058
3	0.250	2.083	0.9483	0.423	1.577	0.7804	0.250	1.577	0.8533
4	0.200	2.283	0.9826	0.500	1.500	0.7581	0.200	1.500	0.8653
5	0.167	2.450	0.9947	0.553	1.447	0.7471	0.167	1.447	0.8616
6	0.143	2.593	0.9989	0.592	1.408	0.7376	0.143	1.408	0.8601
7	0.125	2.718	0.9998	0.622	1.378	0.7189	0.125	1.378	0.8589
8	0.111	2.829	1.0000	0.646	1.354	0.7216	0.111	1.354	0.8585
9	0.100	2.929	1.0000	0.667	1.333	0.7170	0.100	1.333	0.8579

Table 2

The quantity m in the 1st column is the number of decay events, the next are CI: $[u_1, u_m]$ and the corresponding covering probability $pr1$; then the CI $[T - \sigma, T + \sigma]$, and the covering probability $pr2$; then the CI $[u_1, T + \sigma]$, and the covering probability $pr3$.

One sees that the choice of the CI is not simple: in dependence on the preferred criterion they will be different.

References

[1] B.V.Gnedenko, Theory of Probability, English translation, Mir Publishers, 1978.

[2] John E. Freund.
Mathematical statistics. PRENTICE-HALL,INC. Englewood Cliffs, N.J.(1962).

[3] J.Khuyagbaatar, A.Yakushev, Ch.E.Duellmaa, et al.,
$^{48}Ca\ +\ ^{249}Bk$ Fusion Reaction Leading to Element Z=117; Long-Lived α-Decaying ^{270}Db and Discovery of ^{266}Lr
Phys.Rev.Lett.**112** 172501 (2014).

[4] Wilks S.S. Mathematical statistics. Princeton, N.J Princeton University Press, (1947).

[5] Zlokazov V.B.
Confidence Interval Optimization for Testing Hypotheses under Data with Low Statistics.
Comp.Phys.Comm., 185 (2014) pp 933-938.

[6] V.B.Zlokazov Radioactivity. Case: rare events.
Physics of Particles and Nuclei, Letters. 2015 v.12 N2 pp.262-268

[7] V.B. Zlokazov and V.K. Utyonkov Analysis of decay chains of superheavy nuclei produced in the 249Bk+48Ca and 243Am+48Ca reactions J. Phys. G: Nucl. Part. Phys. 44 (2017) 075107 (13pp)