Statistical analysis of rare events—synthesis of the element 114

V.B. Zlokazov
FLNP, JINR, 141980 Dubna, Moscow region, Russia

Received: 24 January 2000 / Revised version: 23 March 2000
Communicated by B. Povh

Abstract. Rare events pose a problem: is an observed chain of radioactive decays that of the background, or are they genetically linked? The paper suggests an approach for the problem solution, based on formalization of the background concept. This approach is an inevitable alternative to other methods, which require the a priori information about the linked decays, in a situation when such information is absent, but, instead, the background information is available, e.g., from the calibration measurements. The method is illustrated by the analysis of data registered in the experiment on the synthesis of the element 114 as one of practically important examples of the analysis of rare events.

The logic and the apparatus of nuclear experiments getting ever more complicated, the situation arises when the result of such experiments is the observation of one single event, which admits a multiple interpretation, and first of all, as a random signal combination.

The direct use of statistical methods in this case is either impossible, or inappropriate: a combination with methods of probability theory is needed. Of course, the mathematical analysis in such situation loses the reliability and safety of the classical statistics; but it allows us to extract the optimum volume of information from the data of a small size which is possible in this case at all.

The observed event is in the best case a sequence of some subevents, or otherwise, signals, which cannot be identified even formally in the sense that they almost all are results of some radioactive decays, statistically independent of each other, and it is very difficult to decide from what decays they come—from those of interest or some others.

So both background signals and those of genetically linked decays of interest are random, statistically independent and formally undistinguishable. The only chance to separate them is given by differences of time characteristics of their combinations, or, speaking more generally, by the differences of probabilistic characteristics (means, variances, frequencies, etc.) of these combinations.

Bearing this in mind, one can build two approaches to tackle with the problem of signal identification:

- Formalize the concept of a background signal combination (BSC) and test whether the signal sequence analyzed does fit in this concept or not.
- Formalize the concept of a linked decay signal combination (LDSC) and test whether the signal sequence analyzed does fit in this concept or not.

The authors of [1] (Dr. K.H. Schmidt and his colleagues) have preferred the latter approach (correlation analysis in their terms). Their method is widely used now, but its success strongly depends on the volume and quality of the a priori information about the qualitative structure of the event (list of classes in terms of [1]) and the half-lives of its constituents.

In cases of extremely indefinite and poor experimental outcomes we do not know the structure of the decay chain a priori; neither the reliable information about the half-lives of members of this chain is available. In this situation the first approach is more attractive: see whether the signal group analyzed corresponds statistically to the pattern of the BSC or not; if not then the next analysis comes trying to find the informative pattern for the event interpretation. The most important advantage of this approach is the fact that for building the BSC pattern we can use the objective sources of information—data of the background calibration measurement, which, in addition, are not affected by poor statistics. The first approach is not competitive against the second one; rather, it is a necessity to which one resorts when there is a complete absence of reliable information about the characteristics of the physical process.

Below some mathematical tools needed for further analysis are described.

1 Functions of probability distribution for the radioactive decay

The classical function of probability distribution for an event (the radioactive decay) at a time moment \( t \) is:

\[
P(t) = 1 - \exp(-lt),
\]

where \( l = \ln(2)/T \), and \( T \) is the half-life of the nucleus. The density of this probability is
and built a quadratic form
\[ \hat{S}_n = \frac{1}{n} \sum_{i,j=1}^{n} c_{ij}^{-1} x_i x_j, \]
it would have the $\chi^2_n$ distribution. But the inversion of a matrix like (5) is rather complicated; besides, the probability of large deviations of a $\chi^2$-distributed random quantity is substantially larger than that of $S_n$ (e.g., for $n = 4$ about 0.07 and 0.045, respectively) and the decision making procedure based on the use of $\hat{S}_n$ loses part of its efficiency specifically in this case.

Thus, preferable is a method based on the direct use of $S_n$. The probability distribution function and its density for (4) can be easily calculated numerically. One can show, that the expectation of (4) is equal to $n$, and the variance to $3n - 1$. Below a table of values of $P(t)$ for $n = 4$ is given.

<table>
<thead>
<tr>
<th>$P$</th>
<th>.05</th>
<th>.10</th>
<th>.15</th>
<th>.20</th>
<th>.25</th>
<th>.30</th>
<th>.35</th>
<th>.40</th>
<th>.45</th>
<th>.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_4$</td>
<td>.6</td>
<td>.9</td>
<td>1.1</td>
<td>1.4</td>
<td>1.6</td>
<td>1.9</td>
<td>2.2</td>
<td>2.5</td>
<td>2.8</td>
<td>3.1</td>
</tr>
<tr>
<td>$P$</td>
<td>.55</td>
<td>.60</td>
<td>.65</td>
<td>.70</td>
<td>.75</td>
<td>.80</td>
<td>.85</td>
<td>.90</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td>3.4</td>
<td>3.8</td>
<td>4.2</td>
<td>4.7</td>
<td>5.3</td>
<td>6.0</td>
<td>6.9</td>
<td>8.2</td>
<td>10.3</td>
<td></td>
</tr>
</tbody>
</table>

Numbers in each upper row are probabilities $P$ with the step 0.05, and in the lower one the corresponding values of $S_4$.

With the help of such table (certainly, more detailed) we can construct the 67% confidence interval of $S_4$ as $n \pm \sigma$; corrected for the asymmetry it is: (1.8–10.8).

### 3 The formalism of stochastic Poisson time processes

These are the time functions $K(t_1, t_2)$—number of random events, occurred during a time interval $(t_1, t_2)$ with a probability $Q_k(t_1, t_2)$ and having the following properties:

1. stationarity: $Q_k(t_1, t_2) = Q_k(t_1 - t_1 - t_2)$ for arbitrary $t_1, t_2$;
2. $Q_k(t_1, t_2)$ independence of the event prehistory: $Q_k(t_1, t_2) = Q_k(t_1, t_2)$, where $C$ means events which happened before $t_1$;
3. rareness of events: $Q_k>1(\delta t) = o(\delta t)$.

These properties allow us to write simply $K(t)$ and $Q_k(t)$ bearing in mind that $t$ means the duration of the time interval considered.

The Poisson processes play an important role in analytical modelling of the stochastic time event background in scientific and technical applications because random events very often satisfy the above-numbered requirements. They disable the need to use the computer simulation to get the estimates of the random background characteristics.

The function of probability distribution of $K(t)$ is
\[ Q_k(t) = \frac{(lt)^k}{k!} \exp(-lt), \]