3-1-2022

How to Describe Hypothetic Truly Rare Events (with Probability 0)

Luc Longpre
The University of Texas at El Paso, longpre@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-22-38

How to Describe Hypothetic Truly Rare Events (with Probability 0)

Luc Longpré and Vladik Kreinovich

Abstract In probability theory, rare events are usually described as events with low probability p, i.e., events for which in N observations, the event happens $n(N) \sim p \cdot N$ times. Physicists and philosophers suggested that there may be events which are even rarer, in which $n(N)$ grows slower than N. However, this idea has not been developed, since it was not clear how to describe it in precise terms. In this paper, we propose a possible precise description of this idea, and we use this description to answer a natural question: when two different functions $n(N)$ lead to the same class of possible "truly rare" sequences.

1 Formulation of the Problem

How rare events are described now. The current approach to describing rare events comes from probability theory.

In the probability theory, each class of events is characterized by its probability p. From the observational viewpoint, probability means a frequency with which such an event is observed in a long series of observations. Specifically, this means that if we make a large number N of observations, then the number of observations $n(N)$ in which the event occurred is asymptotically proportional to $n(N) \sim p \cdot N$; see, e.g., [2].

In these terms, rare events correspond to cases when the probability p is small: the smaller the probability p, the fewer the occurrences of the events, so the rarer is this event.

Challenge. Probability theory deals with events for which the number of occurrences $n(N)$ is proportional to the number of observations N. But what if we have

Luc Longpré and Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, TX 79968, USA, e-mail: longpre@utep.edu, vladik@utep.edu

an event which is even rarer, i.e., an event for which the number of occurrences $n(N)$ grows even slower – e.g., as $C \cdot N^\alpha$ for some $C > 0$ and $\alpha < 1$?

Such a hypothetical possibility was proposed, on the intuitive level, by several physicists and philosophers. This idea may sound interesting, but the big challenge is that it is not clear how to develop it into something quantitative – since, in contrast to probability theory, there is no readily available formalism to describe such rare events.

What we do in this paper. In this paper, we provide a possible way to formally describe such rare events – and thus, to be able to analyze them.

2 Analysis of the Problem and the Resulting Proposal

Analysis of the situation. In general, when the expected number of events observed during the first N-th observations is described by a formula $n(N)$, the expected number of events occurring during the N observations can be estimated as the difference $n(N) - n(N-1)$ between:

- the overall number $n(N)$ of events observed during the first N observations and
- the overall number $n(N-1)$ of events observed during the first $N-1$ observations.

This expected number of events occurring during a single observation is nothing else but the probability p_N that the N-th observation leads to the desired effect.

In the traditional probability approach, when $n(N) = p \cdot N$, this probability is equal to p:

$$p_N = n(N) - n(N-1) = p \cdot N - p \cdot (N-1) = p.$$

In the desired random case, the corresponding difference decreases with N. For example, for $n(N) \sim c \cdot N^\alpha$, we have

$$p_N = C \cdot N^\alpha - c \cdot (N-1)^\alpha = c \cdot N^\alpha \cdot \left(1 - \left(\frac{N-1}{N}\right)^\alpha\right).$$

Here,

$$\left(\frac{N-1}{N}\right)^\alpha = \left(1 - \frac{1}{N}\right)^\alpha \sim 1 - \frac{\alpha}{N},$$

and therefore

$$p_n \sim C \cdot N^\alpha \cdot \frac{\alpha}{N} = c \cdot N^a,$$

where we denoted $c \stackrel{\text{def}}{=} C \cdot \alpha$ and $a \stackrel{\text{def}}{=} \alpha - 1$.

This leads to the following proposal.

Proposal. We describe a rare event with $n(N)/N \to 0$ as a random sequence in which the N-th observation leads to an event with probability $p_N = n(N) - n(N-1)$, and the occurrences for different N are independent.

3 A Natural Question: Which Probabilities p_N and q_N Lead to the Same Class of Rare Observations?

Now we can start analyzing the new notion of rareness. The fact that now we have a formal definition enables us to start analyzing quantitative questions. Let us start with the question described in the title of this section.

Why this question is important. In the usual probabilistic setting, we need only one parameter to describe the degree of rareness: the probability p. The value of this parameter can be determined based on the observations, as the limit of $n(N)/N$ when N increases. The more observations we make, the more accurately we can determine this probability.

In contrast, when we have a sequence of events with different probabilities p_N, for each N, we only have one case in which the event either occurred or did not occur. Clearly, this information is not sufficient to uniquely determine the corresponding probability p_N. In other words, this means that different sequences p_N can lead to the exact same sequence of observations. In mathematical terms, based on the observations, we cannot uniquely determine the sequence p_N, we can only determine a class of sequences consistent with this observation.

This leads to the following natural questions: what is this class? Or, in other terms: when do two sequences p_N and q_N lead to the same sequence of observations?

Let us describe this question in precise terms. The usual way to apply probabilistic knowledge to practice is as follows: we prove that some formula is true with probability 1, and we conclude that this formula must be true for the actual random sequence.

For example, it is proven than if we have a sequence of independent events with probability p each, then the ratio $\widetilde{n}(N)/N$, where $\widetilde{n}(N)$ denotes the actual number of occurrences of this event in the first N observations, tends to p with probability 1. We therefore conclude that for the actual sequence of observations, we will have $\widetilde{n}(N)/N \to p$. Similarly, we conclude that the actual sequence of observations should satisfy the Central Limit Theorem, according to which the distribution of the deviations $\widetilde{n}(N)/n - p$ of this ratio from the actual probability is asymptotically normal, etc.

In other words, the usual application assumes that when we have a random sequence, it should satisfy all the laws that occur with probability 1 – or, equivalently, that this sequence should not belong to any definable set of probability measure 0. This idea underlies the formal definition of a random sequence in Algorithmic Information Theory – it is defined as a sequence that does not belong to any definable set of measure 0; see, e.g., [1].

Here, "definable" can be formalized in different ways: computable in some sense, describable by a formula from some class, etc. No matter how we describe it, there are only countably many such sets. So, the measure of their union is 0, and if we delete this union from the set of all possible binary sequences, we therefore still get a set of measure 1.

Let us apply this formalization to our question. When we select the probabilities p_N, we thus determine a probability measure on the set of all binary sequences, and we can use the above definition to formally define when a binary sequence is random with respect to this measure. So, the above question takes the following form: for which pairs of sequences p_N and q_N there exists a binary sequence which is random with respect to both measures? Let us call such probabilities p_N and q_n *indistinguishable*.

The following proposition explains when two sequences are indistinguishable.

Proposition. *Sequences $p_N \to 0$ and $q_N \to 0$ are indistinguishable if and only if*

$$\sum_{N=1}^{\infty} (\sqrt{p_N} - \sqrt{q_N})^2 < \infty.$$

Proof. According to [3] (see also Problem 4.5.14 in [1]), the two probabilities p_N and q_N are indistinguishable if and only if

$$\sum_{N=1}^{\infty} \left[(\sqrt{p_N} - \sqrt{q_N})^2 + \left(\sqrt{1-p_N} - \sqrt{1-q_N} \right)^2 \right] < \infty. \quad (1)$$

In our case, when $p_N \to 0$ and $q_N \to 0$, we have $\sqrt{1-p_N} \sim 1 - 0.5p_N$, thus

$$\sqrt{1-p_N} - \sqrt{1-q_N} \sim -0.5 \cdot (p_N - q_N) = -0.5 \cdot (\sqrt{p_N} - \sqrt{q_N}) \cdot (\sqrt{p_N} + \sqrt{q_N}).$$

Therefore,

$$\left(\sqrt{1-p_N} - \sqrt{1-q_N} \right)^2 \sim 0.25 \cdot (\sqrt{p_N} - \sqrt{q_N})^2 \cdot (\sqrt{p_N} + \sqrt{q_N})^2.$$

Here, $p_N \to 0$ and $q_N \to 0$, thus $\left(\sqrt{p_N} + \sqrt{q_N}\right)^2 \to 0$. Hence, the second term in the left-hand side of the formula (1) is asymptotically negligible in comparison with the first term, and the convergence of the sum is thus equivalent to the convergence of the first terms.

The proposition is proven.

Comment. By using this proposition, one can check that the probabilities

$$p_N = c \cdot N^a \text{ and } q_N = d \cdot N^b$$

are indistinguishable only when they coincide. Thus, all the probabilities $p_N = c \cdot N^a$ corresponding to different pairs (c, a) are distinguishable.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. M. Li and P. Vitanyi, *An Introduction of Kolmogorov Complexity and Its Applications*, Springer Verlag, New York, 2008.
2. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011
3. V. G. Vovk, “On a randomness criterion”, *Soviet Math. Doklady*, 1987, Vol. 35, pp. 656–660.