Fuzzy Techniques Explain Empirical Power Law Governing Wars and Terrorist Attacks

Hung T. Nguyen
Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA, and
Faculty of Economics, Chiang Mai University
Chiang Mai, Thailand
hunguyen@nmsu.edu

Kittawit Autchariyapanitkul
Faculty of Economics
Maejo University
Chiang Mai, Thailand
kittar3@hotmail.com

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
vladik@utep.edu

Abstract—The empirical distribution of the number of casualties in wars and terrorist attacks follows a power law with exponent 2.5. So far, there has not been a convincing explanation for this empirical fact. In this paper, we show that by using fuzzy techniques, we can explain this exponent. Interesting, we can also get a similar explanation if we use probabilistic techniques. The fact that two different techniques lead to the same explanation makes us reasonably confident that this explanation is correct.

I. EMPIRICAL POWER LAW GOVERNING WARS AND TERRORIST ATTACKS: INTRODUCTION

For historic data about wars, it is possible to perform statistically reliable analysis. Wars are horrible, people die. Good news is that conflicts with numerous casualties are relatively rare. Bad news that in the history of mankind, there have been so many conflicts with numerous causalities that it is already possible to make statistically significant conclusions based on the historical data.

Empirical power law: a description. One statistical conclusion that researchers have extracted from the analysis of historical data is that the wars follow the power law: the number of conflicts with n casualties decreases with n as $c \cdot n^{-2.5}$; see, e.g., [5].

In other world, the probability density function $\rho(n)$ that describes the distribution of different wars by the number of casualties has, for sufficiently large n (i.e., for $n \geq n_0$ for some n_0), the form

$$\rho(n) = c \cdot n^{-2.5}.$$

Terrorist attacks follow the same pattern. The original power law was used to describe wars, i.e., events that caused at least 1000 casualties. Later, it turned out that the same power law also accurately describes events with fewer casualties, such as terrorist attacks; see, e.g., [2] and references therein.

But why? While the newly arriving data continues to confirm the power law for wars and terrorist attacks, there is still no convincing theoretical explanation for the specific exponent 2.5 of the corresponding power law.

What we do in this paper. In this paper, we show that the use of fuzzy techniques can provide an explanation for the observed exponent of the corresponding power law. We also show that a similar explanation can also be obtained if, to describe the corresponding uncertainty, we consider probabilistic techniques instead of the fuzzy ones.

The fact that two different types of techniques leads to the same explanation makes us more confident that this explanation is correct.

II. GENERAL ANALYSIS OF THE PROBLEM

Let us consider reasonable constraints on the probability density $\rho(n)$. Our objective is to explain why, of all possible power laws $\rho(n) \sim n^{-\alpha}$, wars and terrorist attacks are described by the exponent $\alpha = 2.5$.

To come up with such an explanation, let us first analyze, based on the first principles, which values of the exponent α are reasonable in the description of wars and terrorist attacks.

First idea: the overall expected number of casualties is finite. This sounds trivial from a common sense viewpoint, but this is a serious requirement from a mathematical viewpoint, since for many power laws, the expected value of the corresponding random variable is infinite.

Let us consider this requirement from a mathematical viewpoint. The expected value of the random variable n under the probability distribution described by the probability density function $\rho(n)$ is equal to the integral

$$\int n \cdot \rho(n) \, dn.$$

For the power law $\rho(n) \sim n^{-\alpha}$, the above integral becomes proportional to $\int_{n_0}^{\infty} n^{1-\alpha} \, dn$. This integral, in its turn, is proportional to the difference between the value $n^{2-\alpha}$ at the starting point n_0 and at infinity.

For $n = n_0$, the value of $n^{2-\alpha}$ is, of course, finite. So, to make sure that the whole expected value is finite, we need to ensure that the value $n^{2-\alpha}$ becomes finite at infinity. Thus, we cannot have $\alpha < 2$ – else this value will be infinite.

We cannot have $\alpha = 2$ either – because then, the integral describing the mean takes the form $\int_{n_0}^{\infty} n^{-1} \, dn$, which is equal to $\ln(\infty) - \ln(n_0)$ and thus, to infinity.

So, the natural requirement that the overall expected number of casualties is finite implies that $\alpha > 2$.

Second idea: taking into consideration that wars and terrorist attacks are difficult to predict. It is a known commonsense fact that it is very difficult to predict wars or terrorist attacks.

While this statement makes sense from the viewpoint of common sense, from the viewpoint of the corresponding probabilistic model, this may sound somewhat unexpected. Indeed, the whole purpose of statistical data analysis is to make predictions. Yes, sometimes we do not know the probability distribution; in such cases, of course, prediction is difficult. But here, we know exactly the probability distribution – so why is prediction difficult?

The answer lies in the fact that the accuracy of most statistics-based predictions – starting from the simplest possible case of using arithmetic average as an estimate for the mean – is described in terms of the corresponding standard deviation; se, e.g., [6].

For the usual normal distributions, standard deviations are finite (and usually small), so we can have reasonably accurate predictions. However, for many power laws, variance is infinite (hence standard deviation is infinite), so good predictions are not possible. (That the variance is sometimes infinite is not surprising: as we have mentioned earlier, for the power law, even the means are often infinite.)

In these terms, the common sense idea that wars and terrorist attacks are difficult to predict means that for the specific power that describes these phenomena, the variance should be infinite.

Let us describe this requirement in terms of the exponent α. The variance V is equal to $M_2 - \mu^2$, where

$$M_2 \stackrel{\text{def}}{=} \int n^2 \cdot \rho(n) \, dn$$

is the second moment and μ is the mean. Thus, in situations when the mean is finite, infiniteness of the variance is equivalent to infiniteness of the second moment.

For a general power law $\rho(n) \sim n^{-\alpha}$, the second moment is proportional to

$$\int_{n_0}^{\infty} n^2 \cdot n^{-\alpha} \, dn = \int_{n_0}^{\infty} n^{2-\alpha} \, dn = \left. \frac{n^{3-\alpha}}{3-\alpha} \right|_{n_0}^{\infty}.$$

The value $n^{3-\alpha}$ is finite for $n = n_0$, so for the integral to be infinite, this expression need to be infinite when $n \to \infty$. Thus, we must have $\alpha < 3$.

Conclusion. Our analysis shows that the exponent of the corresponding power law must be between 2 and 3.

Remaining question. Which value from the interval $(2, 3)$ should we choose? Or, stated differently: how can we explain that the empirical data is in very good accordance with one specific value from this interval: namely, the value $\alpha = 2.5$?

What we do next. In the following section, we show that the use of fuzzy techniques explains the use of 2.5. In the section after that, we show that probabilistic techniques can also explain the use of $\alpha = 2.5$.

As we have mentioned, the fact that two different techniques lead to the same conclusion make us confident that this conclusion is indeed true.

III. LET US APPLY FUZZY TECHNIQUES

Why fuzzy? At first glance, we have a very precise description of the situation: we have a value α in an interval $(2, 3)$. This is a very crisp statement, with no uncertainty, no fuzziness.

However, the situation is not as crisp as it may seem at first glance. From the purely mathematical viewpoint, finiteness of the overall losses means that $\alpha > 2$. However, as the value α gets closer and closer to 2, the expected value becomes larger and larger – and thus, unrealistic. In other words, it is not enough to require that the expected number of losses is finite, we also must require that this expected value is not too large – and thus, that α should be significantly larger than 2.

Here comes fuzziness: "not too large" is not a precise term, as well as "significantly larger". To describe such terms, fuzzy logic techniques have been invented [7]; see, e.g., [3], [4].

Similarly, instead of a seemingly crisp inequality $\alpha < 3$, we have, in reality, a fuzzy inequality.

How to describe this fuzziness in precise terms. We want to make sure that both differences $\alpha - 2$ and $3 - \alpha$ are positive – in some fuzzy sense.

Let $\mu(x)$ be a membership function that describes this "positiveness". The larger the positive number, the more confident we are that this number is common-sense positive. Thus, the function $\mu(x)$ should be increasing with x.

In terms of this membership function, for each α, the degree to which the first inequality is satisfied is equal to $\mu(\alpha - 2)$, and the degree to which the second inequality is satisfied is equal to $\mu(3 - \alpha)$.

To find the degree to which both conditions are satisfied, we must apply an appropriate "and"-operation (t-norm) $f_\&(a, b)$ to these two numbers $a = \mu(\alpha - 2)$ and $b = \mu(3 - \alpha)$ and thus, get the value

$$f_\&(\mu(\alpha - 2), \mu(3 - \alpha)).$$

Based on the above precise description of fuzziness, which value α should we select? The above expression describes the degree to which a given value α satisfies both inequalities. If we want to select a single number, a reasonable idea is to select the value α for which we are most confident that this value satisfies both inequalities.

We therefore need to select a value α for which the above expression attains its largest possible value.

From the idea to the actual selection. The resulting selection of α, in general, depends of the choice of a membership function $\mu(x)$ and of the "and"-operation $f_\&(a, b)$.

Since we have no a priori reason to select one or another "and"-operations, let us select the computationally simplest one $f_\&(a, b) = \min(a, b)$. In this case, the optimized function has the form

$$\min(\mu(\alpha - 2), \mu(3 - \alpha)).$$

Let us show that this expression attains its maximum when $\alpha = 2.5$. Indeed, for this particular value α, the above expression takes the form

$$\min(\mu(0.5), \mu(0.5)) = \mu(0.5).$$

On the other hand, for any value $\alpha < 2.5$, we have

$$\alpha - 2 < 0.5 < 3 - \alpha.$$

Thus, due to monotonicity of the membership function $\mu(x)$, we have

$$\mu(\alpha - 2) < \mu(0.5) < \mu(3 - \alpha)$$

and thus,

$$\min(\mu(\alpha - 2), \mu(3 - \alpha)) = \mu(\alpha - 2) < \mu0.5).$$

Similarly, for any value $\alpha > 2.5$, we have

$$3 - \alpha < 0.5 < \alpha - 2.$$

Thus, due to monotonicity of the membership function $\mu(x)$, we have

$$\mu(3 - \alpha) < \mu(0.5) < \mu(\alpha - 2)$$

and thus,

$$\min(\mu(\alpha - 2), \mu(3 - \alpha)) = \mu(3 - \alpha) < \mu(0.5).$$

So, the maximum is indeed attained only when we have $\alpha = 2.5$.

Conclusion. Thus, fuzzy idea indeed explain why we should select the exponent $\alpha = 2.5$ in the description of statistics of wars and terrorist attacks.

IV. LET US APPLY PROBABILISTIC TECHNIQUES

Analysis of the problem: let us introduce a probability distribution on the interval $(2, 3)$**.** We are uncertain about the value α. All we know about the unknown exponent value α is that this value lines somewhere on the interval $(2, 3)$.

In the probabilistic approach, to describe this uncertainty, we can use a probability distribution $f(\alpha)$ on the interval $(2, 3)$; see, e.g., [6]. Of course, there are many different probability distributions on this interval. Which distribution should we choose?

Let us use Lalpace's Indeterminacy Principle. We do not have any reason to believe that some values from this interval are more probable than others. Thus, it make sense to assume that all the values from the interval are equally probable, i.e., that we have a uniform distribution on this interval. This argument – dating back to Laplace, one of the founders of the probability theory – is known as *Laplace's Indeterminacy Principle*.

Its modern form is known as the *Maximum Entropy approach* (see, e.g., [1]) when out of all possible probability distributions, we select a one for which the *entropy*

$$S \stackrel{\text{def}}{=} - \int f(\alpha) \cdot \ln(f(\alpha)) \, d\alpha$$

attains its largest possible value.

For distributions located on the interval $(2, 3)$, the only constraint on the value $f(\alpha)$ is that the overall probability to be on this interval should be equal to 1: $\int_2^3 f(\alpha) \, d\alpha = 1$.

Thus, we arrive at the following constraint optimization problem: maximize the expression $- \int_2^3 f(\alpha) \cdot \ln(f(\alpha)) \, d\alpha$ under the constraint $\int_2^3 f(\alpha) \, d\alpha = 1$. By using the Lagrange multiplier method, we can reduce this constraint optimization problem to the following unconstraint one: maximize

$$- \int_2^3 f(\alpha) \cdot \ln(f(\alpha)) \, d\alpha + \lambda \cdot \left(\int_2^3 f(\alpha) \, d\alpha - 1 \right),$$

where the constant α must be defined from the condition that the constraint $\int_2^3 f(\alpha) \, d\alpha = 1$ be satisfied.

Differentiating this expression with respect to each unknown $f(\alpha)$ and equating the derivative to 0, we conclude that

$$- \ln(f(\alpha)) - 1 + \lambda = 0,$$

i.e., that $\ln(f(\alpha)) = \lambda - 1$ and thus, that $f(\alpha) = \exp(\lambda - 1)$. Thus, $f(\alpha)$ does not depend on α, it is a constant $f(\alpha) = \rho_0$, for some constant ρ_0.

The condition $\int_2^3 f(\alpha) \, d\alpha = 1$ now implies that

$$\int_2^3 \rho_0 \, d\alpha = 1,$$

and thus, that $\rho_0 = 1$. So, we indeed have a uniform distribution on the interval $(2, 3)$.

So what value from the interval $(2, 3)$ **should we choose: a problem.** Now we know the reasonable probability distribution on the set of all possible values α – namely, the uniform distribution, with probability density $f(\alpha) = 1$.

Based on this distribution, which value from the interval $(2, 3)$ should we choose?

How to select a single value: idea. A reasonable idea is to decrease the expected loss caused by an erroneous choice of α. If we denote the loss caused by selecting a value $\widetilde{\alpha}$ when the actual value is α by $L(\widetilde{\alpha}, \alpha)$, then minimizing the expected loss means minimizing the expression $\int L(\widetilde{\alpha}, \alpha) \cdot f(\alpha) \, d\alpha$.

How to describe the corresponding loss function? The loss happens every time the selected value $\widetilde{\alpha}$ is different from the actual value α, i.e., when the difference $\Delta\alpha \stackrel{\text{def}}{=} \widetilde{\alpha} - \alpha$ is different from 0. Thus, it makes sense to consider a loss function which depends on this difference: $L(\widetilde{\alpha}, \alpha) = \ell(\Delta\alpha)$ for an appropriate function $\ell(x)$.

Which function $\ell(x)$ should we choose? It is reasonable to assume that the difference $\Delta\alpha$ is small, so we can expand the dependence $\ell(x)$ in Taylor series and keep the first few terms in this expansion:

$$\ell(x) = \ell_0 + \ell_1 \cdot x + \ell_2 \cdot x^2 + \ldots$$

When we selected α correctly, i.e., when $\Delta\alpha = 0$, we should not have any loss, so we should have $\ell(0) = 0$. Thus, we have $\ell_0 = 0$, and hence,

$$\ell(x) = \ell_1 \cdot x + \ell_2 \cdot x^2.$$

The loss is the smallest when $\Delta\alpha = 0$. So, the function $\ell(x)$ attains its minimum for $x = 0$. Thus, the derivative of the function $\ell(x)$ should be equal to 0 when $x = 0$. This implies that $\ell_1 = 0$. Thus, the loss function is proportional to $(\Delta\alpha)^2$. So, the expected loss is equal to

$$\ell_2 \cdot \int_2^3 (\widetilde{\alpha} - \alpha)^2 \cdot f(\alpha)\, d\alpha.$$

For this loss function, which value α should we choose? Minimizing the expected loss means selecting a value $\widetilde{\alpha}$ for which the above expression attains the smallest possible value. Differentiating this expression with respect to the unknown $\widetilde{\alpha}$ and equating the derivative to 0, we conclude that

$$2\ell_2 \cdot \left(\widetilde{\alpha} - \int_2^3 \alpha \cdot f(\alpha)\, d\alpha \right) = 0,$$

i.e., that'

$$\widetilde{\alpha} = \int_2^3 \alpha \cdot f(\alpha)\, d\alpha.$$

For the uniform distribution $f(\alpha) = 1$, this implies

$$\widetilde{\alpha} = \int_2^3 \alpha\, d\alpha = \left. \frac{\alpha^2}{2} \right|_2^3 = \frac{3^2}{2} - \frac{2^2}{2} = \frac{9-4}{2} = 2.5.$$

Conclusion. In the probabilistic approach, we also arrive at the conclusion that the best value to choose is the value $\alpha = 2.5$ – exactly the same value that we have empirically observed.

References

[1] E. T. Jaynes and G. L. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.
[2] N. F. Johnson et al., "New online ecology of adversarial aggregates: ISIS and beyond", *Science*, 2016, Vol. 352, No. 6292, pp. 1459–1463.
[3] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
[4] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
[5] D. Roberts and D. Turcotte, "Fractality and self-organized criticality of wars", *Fractals: Complex Geometry, Patterns, and Scaling in Nature and Society*, 1998. Vol. 6, No. 4, pp. 351–357.
[6] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.
[7] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.