Fuzzy Techniques Explain Empirical Power Law Governing Wars and Terrorist Attacks

Hung T. Nguyen[1], Kittawit Autchariyapanitkul[2],
and Vladik Kreinovich[3]
[1]Dept. of Mathematical Sciences, New Mexico State University
Las Cruces, NM 88003, USA, and Faculty of Economics
Chiang Mai University, Thailand, hunguyen@nmsu.edu
[2]Faculty of Economics
Maejo University, Chiang Mai, Thailand, kittar3@hotmail.com
[3]Dept. of Computer Science, University of Texas at El Paso
El Paso, Texas 79968, USA, vladik@utep.edu

1. Empirical Law Governing Wars

- A war is defined as a conflict with at least 1000 casualties.
- The number $\rho(n)$ of wars with n casualties is
$$\rho(n) \sim c \cdot n^{-2.5}.$$
- The same power law describes events with fewer casualties, e.g., terrorist attacks.
- Why? There is no convincing theoretical explanation for the specific exponent 2.5.
- In this talk, we show that the use of fuzzy techniques can explain this value.

2. Analysis of the Problem

- Our objective is to explain why, of all possible power laws $\rho(n) \sim n^{-\alpha}$, wars are described by $\alpha = 2.5$.
- First idea – the overall expected number of casualties is finite: $\mu \stackrel{\text{def}}{=} \int n \cdot \rho(n)\, dn < +\infty$.
- For $\rho(n) \sim n^{-\alpha}$, the integral of $n^{1-\alpha}$ is $\sim n^{2-\alpha}$.
- This integral is finite when $\alpha > 2$.
- Second idea: it is very difficult to predict wars or terrorist attacks.
- This statement makes sense from the viewpoint of common sense.
- However, from the viewpoint of the corresponding probabilistic model, this sounds strange.
- Indeed, the whole purpose of statistical data analysis is to make predictions.

3. Analysis of the Problem (cont-d)

- 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
  - is described in terms of the corresponding standard deviation σ.
- This can be shown on the example of using arithmetic average as an estimate for the mean.
- For the usual normal distributions:
  - standard deviations are finite (and usually small),
  - so we can have reasonably accurate predictions.

4. Analysis of the Problem (cont-d)

- However, for many power laws, $V = \infty$ (hence $\sigma = \infty$).
- So, accurate predictions are not possible.
- That V may be infinite is not surprising: for the power law, even the mean μ may be infinite.
- In these terms, the idea that wars are difficult to predict seems to indicate that $V = \infty$.
- $V = M_2 - \mu^2$, where $M_2 \stackrel{\text{def}}{=} \int n^2 \cdot \rho(n)\,dn$.
- Since μ is finite, $V = \infty$ means $M_2 = \infty$.
- For $\rho(n) \sim n^{-\alpha}$, the integral M_2 is $\sim n^{3-\alpha}$.
- So, the integral is infinite when $\alpha < 3$.
- Conclusion: $\alpha \in (2, 3)$.

5. Why Fuzzy?

- At first glance, we have a very crisp description of the situation: $\alpha \in (2, 3)$.
- However, the situation is not as crisp as it may seem.
- As α gets closer and closer to 2, the expected value becomes larger and larger – and thus, unrealistic.
- It is thus not enough to require that the expected number of losses is finite.
- We also must require that this expected value is not too large.
- So, we require that α is significantly larger than 2.
- Here comes fuzziness: "not too large" is not a precise term, as well as "significantly larger".
- Similarly, instead of a seemingly crisp inequality $\alpha < 3$, we have, in reality, a fuzzy inequality.

6. Let Us Use Fuzziness

- 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.
- 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 $\mu(3 - \alpha)$.
- The degree $d(\alpha)$ to which both conditions are satisfied is thus equal to $f_{\&}(\mu(\alpha - 2), \mu(3 - \alpha))$.

7. Using Fuzziness (cont-d)

- It is reasonable to select α for which we are the most confident that both inequalities are satisfied:

$$d(\alpha) \to \max .$$

- We have no a priori reason to select one or another "and"-operations.
- Let us thus et us select the computationally simplest one $f_{\&}(a, b) = \min(a, b)$.
- In this case, $d(\alpha) = \min(\mu(\alpha - 2), \mu(3 - \alpha))$.
- We can prove that this expression attains its maximum when $\alpha = 2.5$.
- Indeed, for this α, $d(\alpha) = \min(\mu(0.5), \mu(0.5)) = \mu(0.5)$.
- On the other hand, for any value $\alpha < 2.5$, we have $\alpha - 2 < 0.5 < 3 - \alpha$.

8. Using Fuzziness: Proof

- For $\alpha < 2.5$, we have $\alpha - 2 < 0.5 < 3 - \alpha$.
- Thus, due to monotonicity of $\mu(x)$, we have $\mu(\alpha - 2) < \mu(0.5) < \mu(3 - \alpha)$ and hence,

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

- Similarly, for any $\alpha > 2.5$, we have $3 - \alpha < 0.5 < \alpha - 2$.
- Thus, due to monotonicity of $\mu(x)$, we have $\mu(3 - \alpha) < \mu(0.5) < \mu(\alpha - 2)$ and hence,

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

- So, the maximum is indeed attained only when we have $\alpha = 2.5$.
- Thus, fuzzy ideas indeed explain why we should select the exponent $\alpha = 2.5$.

9. What If We Apply Probabilistic Techniques?

- All we know about α is that $\alpha \in (2, 3)$.
- To describe this uncertainty, we can use a probability distribution $f(\alpha)$ on the interval $(2, 3)$.
- There are many different probability distributions on this interval.
- Which distribution should we choose?
- 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.
- So, we have a uniform distribution on this interval.
- This argument is known as *Laplace's Indeterminacy Principle*.

10. Probabilistic Approach (cont-d)

- Which value from the interval $(2, 3)$ should we choose?
- A reasonable idea is to decrease the expected loss caused by an erroneous choice of α.
- Let us 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 $\int L(\widetilde{\alpha}, \alpha) \cdot f(\alpha)\, d\alpha$.
- The loss happens every time the selected value $\widetilde{\alpha}$ is different from the actual value α, i.e., when

$$\Delta\alpha \stackrel{\text{def}}{=} \widetilde{\alpha} - \alpha \neq 0.$$

- Thus, it makes sense to consider a loss function which depends on this difference:

$$L(\widetilde{\alpha}, \alpha) = \ell(\Delta\alpha) \text{ for some function } \ell(x).$$

11. Probabilistic Approach (cont-d)

- Which function $\ell(x)$ should we choose?
- It is reasonable to assume that $\Delta\alpha$ 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 $\ell(x) = \ell_1 \cdot x + \ell_2 \cdot x^2$.
- 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 $(\Delta\alpha)^2$.

12. Probabilistic Approach (cont-d)

- So, the expected loss is $\ell_2 \cdot \int_2^3 (\widetilde{\alpha} - \alpha)^2 \cdot f(\alpha)\, d\alpha$.
- Differentiating this expression with respect to $\widetilde{\alpha}$ and equating the derivative to 0, we get

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

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

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

- So, in the probabilistic approach, we also arrive at the conclusion that the best value is $\alpha = 2.5$.
- Two different techniques for describing uncertainty lead to the same explanation.
- This makes us confident that our explanation is correct.

13. Acknowledgments

- This work was supported in part by the National Science Foundation grants:
  - HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and
  - DUE-0926721, and
- by an award from Prudential Foundation.

Home Page

Title Page

◀◀ ▶▶

◀ ▶

Page 14 of 15

Go Back

Full Screen

Close

Quit

14. On Laplace Indeterminacy Principle

- Its modern form is the *Maximum Entropy approach*:
  - out of all possible probability distributions,
  - we select a one for which the *entropy* $S \stackrel{\text{def}}{=} -\int f(\alpha) \cdot \ln(f(\alpha))\,d\alpha$ is the largest possible.
- So, we maximize $-\int_2^3 f(\alpha) \cdot \ln(f(\alpha))\,d\alpha$ under the constraint $\int_2^3 f(\alpha)\,d\alpha = 1$.
- Lagrange multiplier method leas an unconstraint problem: maximize

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

- Differentiating with respect to $f(\alpha)$ and equating the derivative to 0, we get $-\ln(f(\alpha)) - 1 + \lambda = 0$.
- So, $f(\alpha) = \exp(\lambda - 1) = \text{const}$: a uniform distribution.