The Code Underneath

Subjects: Mathematics, Interdisciplinary Applications
Contributor: Julio Rives

The synergy between Newcomb-Benford and Bayes' laws provides a universal framework for comprehending information, probability, conformality, and computational intelligence.

Newcomb-Benford Law | harmt (harmonic unit of information) | likelihood | Canonical PMF | Global-local duality | Bayesian Law | Secretary problem | Cross-ratio | Coding source | Conformability

1. List of MSC (Mathematics Subject Classification) epigraphs

• 03D20 - Recursive functions and relations, subrecursive hierarchies

• 11A63 - Radix representation; digital problems

• 11A67 - Other number representations

• 12F99 - None of the above, but in "Field extensions"

• 30C35 - General theory of conformal mappings

• 30F45 - Conformal metrics (hyperbolic, Poincare, distance functions)

• 33B15 - Gamma, beta and polygamma functions

• 60-08 - Computational methods for probability theory problems

• 60A99 - None of the above, but in "Foundations of probability theory"

• 62C10 - Bayesian problems; characterization of Bayes procedures

• 65E10 - Numerical methods in conformal mappings

• 68P30 - Coding and information theory

• 83-10 - Mathematical modeling for problems pertaining to relativity theory

• 93A13 - Hierarchical systems

• 94A17 - Measures of information, entropy

2. Main Ideas

The manuscript addresses topics in mathematical and computational physics, focusing on the Newcomb-Benford Law (NBL) and its implications. The Newcomb-Benford Law (NBL) describes the frequency of leading digits in many natural data sets, where smaller digits appear more frequently than larger ones. The work seeks to explain the origin of the NBL and its consequences, proposing that an inverse-square law probability mass function (PMF) is at the root of the NBL. Information is conveyed through concepts of harmonic and logarithmic likelihood, and probability is a relative likelihood. The NBL is connected to Bayes' Law, which encodes the strength of the relationship between pairs of normalized items in a specific base or radix. The author proposes conformal coding functions that preserve information and granularity, modeling reality at multiple scales.

Under the tail of the inverse-square law PMF at NBL's root, the "canonical" PMF, the global NBL emerges, while the local NBL is a logarithmic version that emerges from the global one. The proposed PMF is well-defined, with positive probabilities summing to 1, no bias, and is extendable to integers. Global and local Bayesian coding are discussed in terms of how information is processed and represented on harmonic and logarithmic scales. The probability of jumps between consecutive quanta or digits is analyzed, showing that lower energy levels are more stable. The author describes a conformal model that maps external space to a local coding space, preserving information. Hyperbolic distance is used to measure the distance between points in a conformal space, with implications for special relativity.

The NBL reflects the efficiency and economy of nature, where smaller numbers are more accessible at a lower cost. Information is fundamentally rational and relational, with the NBL and Bayes' laws providing a framework for understanding probability and information. The Newcomb-Benford Law and Bayes' Law are fundamental to universal information coding, linking mathematics to physics and providing a basis for understanding reality at multiple scales.

The work explores the Newcomb-Benford Law (NBL) and its manifestations in various fields, highlighting the prevalence of smaller numbers in natural data. The review includes historical and contemporary references, such as the works of Newcomb, Benford, Hill, and others. The methodology adopted departs from a probability mass function (PMF) based on the inverse-square law, arguing that this is the root of the NBL. The approach is mathematical and theoretical, utilizing concepts from information theory, probability, and conformal geometry. The derivation of the global and local NBL from the canonical PMF shows how harmonic and logarithmic information relate to probability. Additionally, the methodology includes the application of Bayes' rule for coding and recoding data, reinforcing the connection between the NBL and information theory. The conformal coding approach and the analysis of elementary jumps and optimal stopping problems are examples of how the methodology is applied to practical and theoretical problems, providing insights into the stability and reactivity of energy levels in physical systems.

3. Coding

Bayesian encoding, recoding, and decoding are elemental computing routines that handle odds. The Bayesian encoding of the relation between two numbers is the entropic allocation of their correlation for a harmonic scale, i.e., their ratio squared multiplied by the probability of the associated interval in the chosen base (see Fig. [fig:InteractingScales] in green). The Bayesian encoding of the relation between two quanta is the entropic contribution of their correlation for a logarithmic scale, i.e., their NBL ratio multiplied by the probability of the associated bucket in the chosen radix (see Fig. [fig:InteractingScales] in blue). This pattern allows us to interpret a Bayesian rule as the formula to encode the rational point n/d or the corresponding range [n,d) of integers; this duality principle asserting that points and lines are interchangeable is endemic to the cosmos.

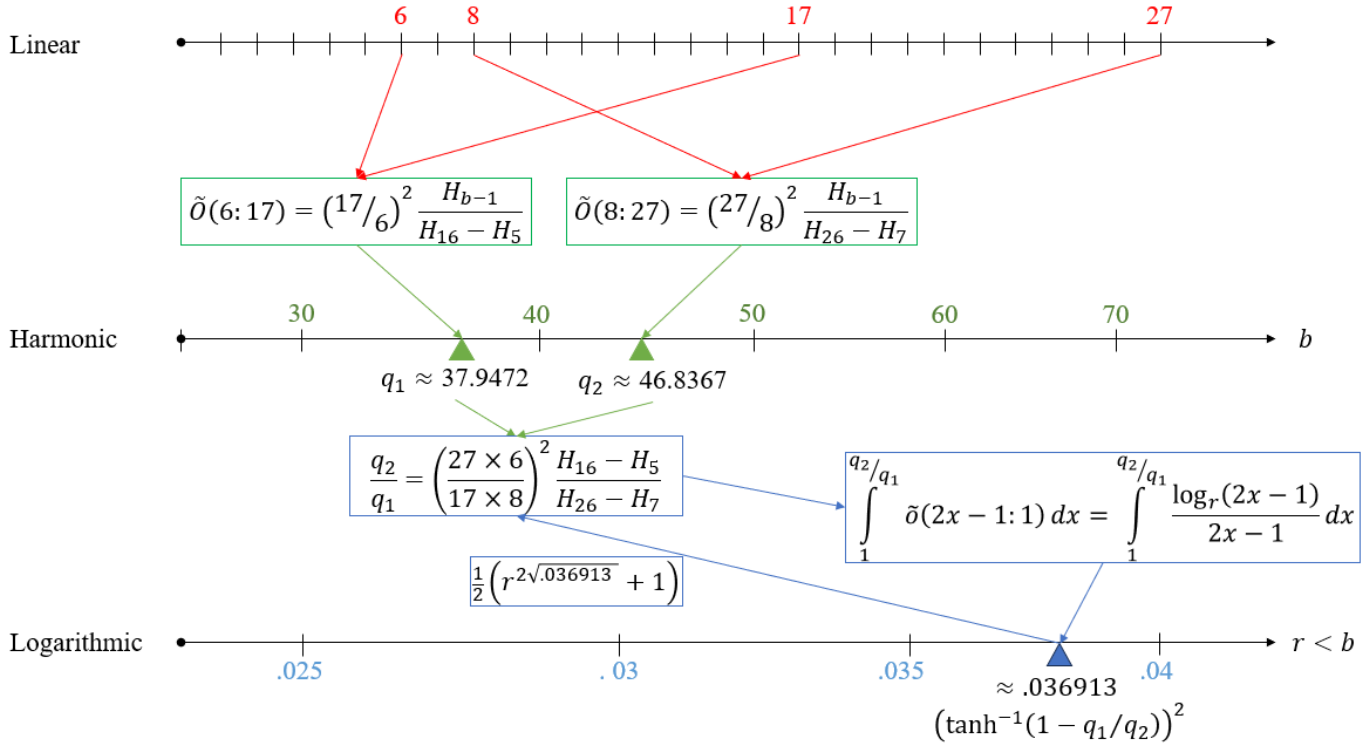

Fig. InteractingScales. Example with b=101 and r=e (Euler's number) of how the three scales interact using the Bayes' rules. q2/q1 is a global Bayesian datum corresponding to the local Bayesian datum Q=1-q1/q2; both represent the Euclidean distance to the origin. Within the logarithmic space of a coding source, the hyperbolic distance and the differential entropy from the origin to Q are 2 artanh(Q) and (artanh(Q))^2=0.036913, respectively. The decoding function returns the latter value to q2/q1. Alternatively, the coding source can directly encode q2/q1 as log_r(2 q2/q1 - 1)=0.384255 and decode it as (1 + r^0.384255)/2.

4. Conformality

The discussion on conformal coding and conformal metrics elucidates how information can be preserved and transformed at different scales. The analysis of optimal stopping problems, such as the secretary problem, illustrates the practical applicability of the developed theory.

5. Conclusions

The research explains how discreteness and the continuum interact under the shelter of the canonical PMF. A complex system and its environment embody the continuous local and discrete global. The NBL probability's derivative of the local takes us to the global, and vice versa; the global's integral situates us in a local setting of likelihood-based probability. A harmonic scale of rational numbers supports the global realm, while a logarithmic scale of "real values" supports the local realm. The harmonic scale's base confines the rational setting, and a logarithmic scale's radix is an exponentiation constant that normalizes a complex system's conformal space of "continuos" information coded in PN.

The work considers a complex system a coding source that observes the outside, operates internally with the gathered information, and takes action on the environment. More precisely, a coding source uses the synergy between Benford's and Bayes's laws to reflect (encode) the external world, process (recode or arithmetically transform) the information, and return (decode) the results to its immediate surroundings.