The Game of Normal Numbers

Dan Gnanapragasam

June 2, 2009

Contents

1 Introduction **2**

2 Normal Numbers and Their Extensions **2**

2.1 Normal Numbers 2

2.2 Extended Normal Numbers 3

3 The Game of Normal Numbers **4**

3.1 Important Terminology 4

3.2 Description of the Game 4

3.3 Payoff Game 5

3.4 Construction of Normal Numbers 5

4 Conclusion **7**

5 Appendix **8**

1 Introduction

Game Theory is a branch of mathematics that uses conditional probability in order to analyze the outcomes of a multi-player situation. Rubenstein (1991) defines a **game** as a "game form" and a "strategy". The game form is the set of rules and procedures that govern the game, as well as the game objectives. The strategy encompasses the processes that the players of the game take in order to complete the game. There are many examples of two-player games in game theory including the famous "prisoner's dilemma" and the whimsical "rock, paper, scissors". Ehud Lehrer's paper analyzes a two-player, zero-sum game in which Player 2 chooses a "distribution" and Player 1 chooses a "realization" and shows that regardless of what Player 2 does, Player 1 can construct a sequence of outcomes that is normal with respect to the sequence of distributions. The rest of this paper discusses the findings of his paper.

2 Normal Numbers and Their Extensions

2.1 Normal Numbers

A **normal number** with respect to a base or expansion is one so that every number, or finite sequence of digits, has an equal probability of occurring. Alternatively, we can loosely state that a number is considered normal with respect to some expansion if the frequency of any finite sequence of n digits has the same chance of appearing in that number as any other sequence of n digits. For example, consider an infinite binary sequence S where S is normal with respect to the binary expansion. Then, for any finite subsequence of zeros and ones, S', contained in the infinite sequence, the relative frequency of either 0s or 1s in S that occur after S' tends to one-half. Other examples of normal numbers (with respect to base 10) include the number $0.1234567891011121314\ldots$ (called the Champernowne constant) and $0.23571113\ldots$ (called the Copeland-Erdös constant). This leads us to the framework for extended normal numbers.

First, we need to formulate a definition of probability in the context of extended normal numbers.

Definition 2.1: *Let X be a finite set of digits and let θ be a probability distribution over $X^{\mathbb{N}}$, the Cartesian product of X with itself countably many times. For any $x^{\mathbb{N}} \in X^{\mathbb{N}}$, and $l \in \mathbb{N}$, with $l < n$, denote $x^{l,n} = (x_l, x_{l+1}, \ldots, x_n)$ and $x^n = x^{1,n}$. The probability with respect to θ that $z \in X$ will appear after x^n is denoted by $\theta(z|x^n)$*

We will refer to instances of $x^{l,n}$ as **strings**. We will later use this to see if certain strings appear within other strings, especially at the end of strings.

Definition 2.2: *Given $k \in \mathbb{N}$, $z^k = (z_1, \ldots, z_k) \in X^k$, $x^{\mathbb{N}} \in X^{\mathbb{N}}$ and $n \in \mathbb{N}$, let $I(x^{n-1}, z^k)$ be 1 if $x^{n-k,n-1} = z^k$ and 0 otherwise.*

This means that $I(x^{n-1}, z^k)$ serves as a parity indicator. If $x^{n-k,n-1} = (x_{n-k}, x_{n-k+1}, \ldots, x^{n-1}) = (z_1, x_2, \ldots, z^k) = z^k$, then $I(x^{n-1}, z^k)$ detects parity and is equal to 1. If $x^{n-k,n-1} = (x_{n-k}, x_{n-k+1}, \ldots, x^{n-1}) \neq (z_1, x_2, \ldots, z^k) = z^k$, then $I(x^{n-1}, z^k)$ detects the inequality and evaluates to 0.

Definition 2.3: *Given* $0 \leq k < n$ *denote* $\bar{I}(x^{n-1}, z^k) = \sum s = k^{n-1} I(x^s, z^k)$. *When* $0 < k$, $\bar{I}(x^{n-1}, z^k)$ *is the number of times the string* z^k *appears in* x^{n-1}. *Let* $z \in X$ *and*

$$Y_\theta(x^n; z^k, z) = \left[\mathbb{I}(x_n = z) - \theta(z|x^{n-1})\right] I(x^{n-1}, z^k),$$

with $\mathbb{I}$ *as the characteristic function.*

As a result of Definition 2.3, Y_θ is restricted to three different values. If x^{n-1} does not end with z^k, then $Y_\theta = 0$; if x^{n-1} ends with z^k and the last digit of x^n is z, then $Y_\theta = 1 - \theta(z|x^{n-1})$; if x^{n-1} ends with z^k and the last digit of x^n is not z, then $Y_\theta = -\theta(z|x^{n-1})$.

Definition 2.4: *Let*

$$\bar{Y}_\theta(x^n; z^k, z) = \frac{\sum_{s=k+1}^{n} Y_\theta(x^s; z^k, z)}{\bar{I}(x^{n-1}, z^k)}$$

and let $\frac{0}{0} = 0$.

This means that $\bar{Y}_\theta$ describes the difference between actual results of the value of z and the average of the probabilities (if we restrict the set that we are analyzing to strings that end with z^k).

2.2 Extended Normal Numbers

Given the definitions listed above, we can now give a definition of a θ-extended normal number.

Definition 2.5 (Extended Normal Numbers): *A sequence* $x^{\mathbb{N}} = (x_1, x_2, \ldots) \in X^{\mathbb{N}}$ *is called* θ*-normal (or a* θ*-extended normal number or normal with respect to* θ*) if*
(a) for any $n = 0, 1, \ldots,$ $\theta(x^n) > 0$
(b) for any $k = 0, 1, \ldots, z^k = (z_1, \ldots, z_k) \in X^k$ *and for every* $x \in X$,

$$\lim_{n\to\infty} \bar{Y}_\theta(x^n; z^k, z) = 0,$$

whenever $\bar{I}(x^n, z^k) \to \infty$ *with* n.

Then, for a sequence to be θ-extended normal it must have the property that any string in the sequence starting at the first index must have a positive probability of occurring [by (a)] and the difference between the empirical results and average probabilities go to 0 as n tends to infinity [by (b)].

3 The Game of Normal Numbers

3.1 Important Terminology

We will define a **zero-sum game** as one where one player's loss is balanced out by another player's gain and vice versa. That is, the gain of any one player, directly results in the loss of some other player, which means that the sum of the gains and losses is always zero. The game that we are dealing with is a two-player game, meaning that that the winning cases are mutually exclusive. Hence, Player 2 wins if and only if Player 1 loses.

A **pure optimal strategy** is a strategy that describes a step-by-step process for one player that guarantees that that player will attain his or her desired result. As such, a random choice system may constitute an optimal strategy, but would not be a pure optimal strategy as it does not, strictly speaking, give a step by step-by-step process for winning the game.

A **convex combination** is a linear combination $a_1x_1 + a_2x_2 + \ldots a_nx_n$ so that the coefficients satisfy $\sum_{j=1}^{n} a_j = -1$ and $a_j \leq 0$ for all $1 \leq j \leq n$

3.2 Description of the Game

Lehrer poses a two-player game that consists of a infinite number of periods. During each period Player 2 chooses a distribution from the set of distributions over X. Then, Player 1 chooses X, a finite set of digits. Player 1's choices depend on his/her previous choices and the previous choices of Player 2 and Player 2 choices depend on his/her previous choices and the previous choices of Player 1. More formally, if we let X be the set of actions of Player 1 and $\Delta(X)$ be the actions of Player 2, then Player's 1 strategy on the n^{th} turn is a function

$$f : \bigcup_{n=0}^{\infty} ((\Delta(X) \times X)^n \times \Delta(X)) \mapsto X$$

and Player 2's strategy on the n^{th} turn is a function

$$\theta : \bigcup_{n=0}^{\infty} ((\Delta(X) \times X)^n \mapsto \Delta(X)$$

We will now use the Kolmogorov extension theorem, which shows that θ creates a probability distribution (also called θ) on $X^{\mathbb{N}}$. Hence, Player 1's turns produce an infinite sequence of choices of X while Player 2's choices produce a distribution θ on $X^{\mathbb{N}}$. Player 1 wins if and only if the sequence of Player 1's turns is θ-extended normal.

3.3 Payoff Game

Let $\Omega = \bigcup_{k=0}^{\infty}(X^k \times X)$ and $\mu(z^k, z) = 2(|X|)^{-(k+1)}$ for $(z^k, z) \in \Omega$. Then, (Ω, μ) is a space of normality tests, that is, the space of tests that must be passed by a number in order to be deemed normal. Define the game $\Gamma^{\mathbb{N}}$ as a game that is the same as the one in the previous section so that the payoff for the game is a random variable over (Ω, μ). Consider stage $n-1$, where Player 1's actions can be described as $x^{n-1} = (x_1, x_2, \dots x_{n-1})$ using the notation introduced in Definition 2.2. Then, define Player 1's actions at the next stage to be x^n and let Player 2's action be $\theta(\cdot|x^{n-1})$; that is the probability of $\cdot$ happening, given that x^{n-1} has happened. Then, the payoff is described by $Y_\theta(x^n; z^k, z)$ at (z^k, z) from Definition 2.3. Moreover, $\bar{Y}_\theta(x^n; z^k, z)$ (from Definition 2.4) is the average random variable payoff on x^n. Then, Player 1 wins $\Gamma^{\mathbb{N}}$ if and only if $\bar{Y}_\theta(x^n; z^k, z) \to 0$ and, by Definition 2.5, if for any $n = 0, 1, \dots, \theta(x^n) > 0$.

3.4 Construction of Normal Numbers

This brings us to the crux of the paper, that Player 1 has a winning strategy for the game of normal numbers.

Theorem 3.1 *Player 1 has a strategy such that for any strategy θ of Player 2, the sequence of Player 1's actions, $(x_1, x_2, \dots)$ is a θ-normal number. Moreover,*
(a) the only information about θ needed to generate the n^{th} digit, x_n is $\theta(\cdot|x_1, \dots x_s)$, $s < n-1$, and the support of $\theta(\cdot|x_1, \dots, x_{n-1})$; and
(b) the algorithm is quadratic (i.e., the number of calculations it requires to compute the n^{th} digit is $O(n^2)$).

Proof. Lehrer's proof produces an inductive strategy for Player 1. Pick $x_1 \in X$ so that $\theta(x_1) > 0$ and suppose that $x^n = (x_1, x_2, \dots, x_n)$ (the strategy of Player 1) has been chosen so that $\theta(x^n) = \theta(x_1, x_2, \dots, x_n) > 0$. Let $\hat{X} = \{z \in X; \theta(z|x_1, x_2, \dots, x_n) > 0\}$ and let $m = |\hat{X}|$. Then, for every $z^k \in X^k$ and $z \in \hat{X}$, let $A(z^k, z)$ be the $m \times m$ matrix so that for $x', x'' \in \hat{X}$,

$$a_{x',x''} = I(x^n, z^k)\left[\mathbb{I}(x' = z) - I(x'' = z)\right].$$

Hence, $A(z^k, z) = 0$ (that is, the zero matrix) if x^n does not end with z^k. If x^n ends with z^k, then row z is identically equal to 1 (except for the diagonal entry) and column z is identically equal to -1 (excluding the diagonal entry). Then, every other element in A is 0.
Let $A(z^k, z)_{x'}$ be row x' of $A(z^k, z)$. Then, $A(z^k, z)$ is defined so that for every pair (z^k, z),

$$Y_\theta((x^n, x'); z^k, z) =< A(z^k, z)_{x'}, \theta(\cdot|x_1, x_2, \dots x_n) >,$$

where $< a, b >$ is the inner product.

Let $\bar{A}$ be the matrix so that:

$$\bar{A} = \sum_{(z^k,z)\in\Omega \text{ and } z\in\hat{X}} \mu(z^k,z)\frac{\bar{Y}_\theta(x^n;z^k,z)}{\bar{I}(x^n,z^k)}A(z^k,z).$$

Then, $\bar{A}$ is a linear combination of $A(z^k,z)$. As $A(z^k,z) = 0$ if x^n does not end with z^k, we know that there are n different matrices $A(z^k,z)$ that are not identically 0. As $A(z^k,z)$ is of the form $(b_i - b_j)_{ij}$ for some $(b_1,\dots,b_m) \in \mathbb{R}^n$, then $\bar{A}$ is also of the form $(b_i - b_j)_{ij}$ for some $(b_1,\dots,b_m) \in \mathbb{R}^n$ as it is a linear combination of matrices of the form of matrix A. Therefore, there exists a row R of $\bar{A}$ so that every element in that row is less than or equal to 0. However, $\bar{A}$ is a zero-sum game, so Player 1 has a pure optimal strategy. Then, let Player 1's action at period $n+1$ be x_{n+1} and let $x_{n+1} \in \hat{X}$ be in R. Hence $\bar{a}_{x_{n+1},x''} \leq 0$ for all x'' in the domain of definition. As a result, $\theta(x^n, x_{n+1}) > 0$ by the definition of $\hat{X}$. As such, any convex combination of the elements of R must also be less than or equal to 0. Then, $\theta(x''|x_1,x_2,\dots,x_n) < 0$. We can state this as

$$\sum_{x''\in\hat{X}} \bar{a}_{x_{n+1},x''}\theta(x''|x_1,x_2,\dots,x_n)$$

$$= \sum_{x''\in\hat{X}}\left[\sum_{(z^k,z)\in\Omega}\mu(z^k,z)\frac{\bar{Y}_\theta(x^n;z^k,z)}{\bar{I}(x^n,z^k)}a_{x_{n+1}}(z^k,z)\right]\theta(x''|x_1,x_2,\dots,x_n) \leq 0,$$

where we define $a_{x_{n+1}}(z^k,z)$ as entry $a_{x_{n+1},x''}$ of matrix $A(z^k,z)$. But as

$$Y_\theta((x^n,x');z^k,z) =< A(z^k,z)_{x'}, \theta(\cdot|x_1,x_2,\dots x_n) >,$$

we can write

$$\sum_{x''\in\hat{X}}\left[\sum_{(z^k,z)\in\Omega}\mu(z^k,z)\frac{\bar{Y}_\theta(x^n;z^k,z)}{\bar{I}(x^n,z^k)}a_{x_{n+1}}(z^k,z)\right]\theta(x''|x_1,x_2,\dots,x_n) \leq 0,$$

as

$$\sum_{(z^k,z)\in\Omega}\mu(z^k,z)\frac{\bar{Y}_\theta(x^n;z^k,z)}{\bar{I}(x^n,z^k)}a_{x_{n+1}}(z^k,z)Y_\theta(z^k,z,z(x^n,x_n+1) \leq 0.$$

At this point, we can simplify notation as follows:

$$Y^{n+1} = Y_\theta((x^n,x_{n+1});z^k,z)$$
$$\bar{Y}^n = \bar{Y}_\theta(x^n;z^k,z)$$
$$\bar{I}^n = \bar{I}(x^n,z^k)$$

Hence, if we defined E(x) to be the expected value of x, we see that

$$E\left(\frac{\bar{Y}^n}{\bar{I}^n}Y^{n+1}\right) \leq 0.$$

Letting n take positive integer values yields a sequence $Y^1, Y^2, \ldots$, and $\bar{I}^1, \bar{I}^2, \ldots$, on Ω. Therefore,
(1) If we let $\bar{I}^0 = 0$, then $\bar{I}^n - \bar{I}^{n-1}$ takes on the values 0 or 1. Also $Y^{n+1} = 0$ if $\bar{I}^n - \bar{I}^{n-1}$.
(2) The Y^n are uniformly bounded.
(3) $\bar{Y}^{n+1} = (\bar{I}^{n-1}\bar{Y}^n + Y^{n+1})/\bar{I}^n$
(4) $E\left(\frac{\bar{Y}^n}{\bar{I}^n} Y^{n+1}\right) \leq 0$.
At this point, we can use Theorem 1 of Lehrer (2002) (see Appendix) and see that $\bar{Y}^n \to 0$ as $\bar{I} \to \infty$. Therefore, the sequence of choices of Player 1 satisfies the conditions of θ-normal, meaning that the target was achieved and Player 1 has a pure optimal strategy. □

Lehrer also goes on to prove that we require $O(n^2)$ operations in order to compute the n^{th} digit of this strategy, meaning that the entire strategy runs in $O(n^3)$ time.

Proof. Consider the matrix $A(z^k, z)$ defined in the proof of part (a) of theorem 1. The number of matrices so that $A(z^k, z) \neq 0$ is $O(n)$ where $A(z^k, z)$ is defined in the n^{th} step of the inductive process. Also, Y_θ^{n+1}, $\overline{I}^n$, and $\overline{Y}_\theta^{n+1}$ are computed in $O(n)$ time. As these processes are nested, the algorithm for picking the n^{th} digit of the sequence is $O(n^2)$, which in turn implies that the algorithm to pick the entire sequence upto the n^{th} digit is $O(n^3)$. □

4 Conclusion

Lehrer's paper serves to illustrate how one can use mathematics to derive order and patterns from a seemingly chaotic system. Though normal numbers are a mathematical construct, their importance is far reaching. Within the field of computer science, normal numbers are truly remarkable as they provide a good definition of what it truly random. By definition, we know that a normal number has the property that any subsequence of length n digits within the normal number has (asymptotically) the same chance of occurring as any other subsequence of digits of length n. In this sense, normal numbers are the ultimate random number generator, better than any pseudo-random generator built into a computer. While it is well-known that almost all real numbers are normal, a proof like Lehrer's that presents an algorithm that shows how to make the number is rather unusual and intriguing.

5 Appendix

The term **almost surely** is commonly used in probability to refer to something that occurs with probability that tends to 1 as the cardinality of the space we are considering increases. For example, in thermodynamics and statistical mechanics considering Einstein solids, energy travels almost surely towards the macropartitions where there are the most total microstates. Understanding the term *almost surely* is important as it plays a part in the following theorem from Lehrer (2002). This theorem is vital for the proof of the theorem shown in Lehrer (2004).

Theorem 1 of Lehrer (2002) : Suppose that
(a) $\{\tilde{\chi}\}_0^\infty$ is a sequence of non-decreasing random variables that assume integer values, $\tilde{\chi}_n - \tilde{\chi}_{n-1} \leq 1$, $\tilde{\chi} \to \infty$ $\mu-$almost surely and $\tilde{\chi}_0 = 0$;
(b) $\{g_n\}$ is an L_2-bounded sequence of random variables in $L_2(\Omega, \mu, \mathscr{F})$ that satisfies $g_n = 0$ whenever $\tilde{\chi}_n - \tilde{\chi}_{n-1} = 0$;
(c) $f_n = \frac{\tilde{\chi}_{n-1} f_{n-1} + g_n}{\tilde{\chi}_n}$, where $f_1 = g_1$; and
(d) $\sum \left\langle \frac{f_{n-1}}{\tilde{\chi}_n}, g_n \right\rangle < \infty$
Then, f_n converges μ-almost surely to zero.

References

[1] Lehrer, E. 2002. Approachability in Infinite Dimensional Space. *International Journal of Game Theory.* **31** 253-268.

[2] Lehrer, E. 2004. The Game of Normal Numbers. *Mathematics of Operations Research.* **29** 259-265.

[3] Rubinstein, A. 1991. Comment on the Interpretation of Game Theory.*Econometrica.* **59** 909-924.