INFORMATION-THEORETIC TECHNIQUE IN NORMAL-NUMBERS

BY

S K. MUKHERJEE

ABSTRACT

In this paper information— theoretic methods have been applied to prove that normality of a real remains invariant under the operation of addition with a real whose entropy lies in a neighbourhood of Zero. This problem has been used to prove that set of reals derived from digits of a normal number by method of difference sequence, are also normal to same scale and conversely. These have been generalised for the K— tuples of numbers.

INTRODUCTION

Shannon (1948) defined entropy in terms of probability of events without restricting the notion of events in any way. Hence, after Shannon, entropy can be defined for a real Dutta and Sen (1967) associated the definition of normality of a real with the entropy of that real. The concept of entropy in normal numbers helps to prove some well-known but complicated results on normal numbers in easier way. Proposition I of this paper is due to Lupskarz & Maxfield [4].

Definition 1. Let \( a \geq 2 \) denote a fixed integer and take \( G \) as the finite set

\[ G = \{0, 1, \ldots, a - 1\} \]

Let further \( K \) denote additive group of reals modulo 1, (0 and 1 identified). Each number \( \xi \) in \( K \) has a unique expansion

\[ \xi = \sum_{k=0}^{\infty} x_k(\xi)a^{-k-1} \pmod{1} \]

with \( x_k(\xi) \in G \), provided we do not allow (as we will) an expansion with \( i_k(\xi) = a - 1 \) for all large \( K \).
The number \( \xi \) is said to be normal to base 'a' if, for each choice of positive integer \( r \) and elements \( y_1, \ldots, y_{r-1} \) in \( G \), one has

\[
\lim_{n \to \infty} \frac{1}{n} \text{number of } 0 < K < n; (x_K(\xi), \ldots, x_{K+r-1}(\xi))=(y_0, \ldots, y_{r-1}) = a^{-r}
\]

In particular, when \( r = 1 \), the number is Simply — normal number.

In literature, normality is defined in two different ways as follows:

**Definition A** A real \( x \) is normal in Scale \( r \) if \( x, rx, r^2x, \ldots \) are Simply — normal in all the scales \( r, r^2, r^3, \ldots \).

**Definition B** A real \( x \) is normal in Scale \( r \) if it is Simply — normal in all the scales \( r, r^2, r^3, \ldots \).

**Definition 2** From Information theory, the entropy of a complete set of events

\[
A_1, A_2, \ldots, A_n
\]

together with their probabilities \( p_1, p_2, \ldots, p_n \) is given by

\[
H(p_1, p_2, \ldots, p_n) = -\sum_{k=1}^{n} p_k \log p_k
\]

the base of logarithm is arbitrary but fixed.

As occurrence of a particular digit may be looked upon as independent random event, so if \( p_i \) is probability of occurrence of \( i^{th} \) digit in \( r — a^n \) representation of \( x \), then

\[
H(x) = -\sum_i p_i \log p_i
\]

According to Dutta \& Sen (1967):

The entropy of a real is maximum if and only if it is Simply — normal number.

If we take scale of representation as base of logarithm, then maximum value of entropy is 1. The property of normality is associated with the invariance of this property of value of entropy being 1 under set of transformation of Scale of notation \( r \) in the set \( \{r, \ r=1, 2, \ldots \} \).

**Definition 3** If \( A \) is input alphabet, \( \mu \) is probability measure defined over Borel field \( F_A \), the stochastic process \([A', F_A, \mu]\) is an information source, where \( A' \) denote class of infinite sequences

\[
x = (\ldots, x_{-1}, x_0, x_1, \ldots)
\]

each \( x_i \) belongs to \( A \).

Since \( A' \) space is fixed by alphabet, Borel field \( F_A \) is always determined by base sets, we can specify a source by \([A, \mu]\). If \( \mu(TS) = \mu(T) \) for any set \( S \in F_A \), the source is stationary, \( T \) being shift — operator.

**Definition 4** A channel is characterized by two alphabets \( A \) and \( B \) and a list of probability measures \( \nu_0 \) defined over Borel field \( F_B \), one for each \( \theta \in A' \) where \( A' \) denote class of infinite sequences

\[
x = (\ldots, x_{-1}, x_0, x_1, \ldots)
\]

each \( x_i \) belongs to \( A \).
We can specify a channel by Symbol \([A, \nu, B]\); we call the channel \([A, \nu, B]\) stationary if, for all \(x \in A\) and \(S \in \mathbb{F}_r\),
\[
\nu_{x+S}(TS) = \nu_{x}(S)
\]
\(T\) is shift — operator.
If every member of input alphabet gives at the output a unique member, then the channel is called a noise-less channel.

**Proposition 1** Let \(C\) be a given digit (or finite sequence of length \(S\) of digits) to scale \(r\). Let \(\lambda\) be a number to scale \(r\) such that number of non-
\(- e \) — digits in first \(N\) digits (or \(SN\) digits) of \(\lambda\) is \(f(N) = o(N)\). We define \(\mathcal{E}\) to be class of elements of type \(\lambda\).

To investigate the entropy of this class \(\mathcal{E}\) and then to prove that when elements of \(\mathcal{E}\) are added to normal numbers a new set of normal numbers is generated.

(The Second part of this problem is due to LUPSKEB & MAXFIELD. here we prove this by applications of Information — Theory)

**Proof** By given hypothesis, number of non-C-digits in first \(N\) digits of \(\lambda\) is \(f(N) = o(N)\). Hence,
\[
\lim_{x \to \infty} \frac{f(N)}{N} = 0
\]
If \(P_C\) be probability of non-C-digits, then
\[
P_C \to 0 \quad \text{as} \quad N \to \infty
\]
Then, as a consequence of Cauchy's first limit theorem, probability of all such non — C — digits tends to Zero and as probability is non-negative, so all such digits have their probabilities lying between 0 and \(\varepsilon\) where \(\varepsilon > 0\) is arbitrary small.

Also, probability of C-digit lies between \(1 - \varepsilon\) and 1.
If we designate by \(H_N\), the entropy of class \(\mathcal{E}\), then \(H_N\) lies in a very close — neighbourhood of 0 and so \(\mathcal{E}\) is a class of almost — completely — non — normal number.

Next, let \(A\) be set of digits obtained by \(r\) — adic representation of normal number \(\alpha\). Then, according to [5], there is a transformation \(T\), defined on unit — interval by \(Tx = rx \pmod{1}\) and there exists a \(T\)-invariant measure \(\mu\), equivalent with Lebesgue — measure and also this \(T\) is ergodic. Thus, considering the above set \(A\) as alphabet, we get a stationary source \([A, \mu]\).

Let \(\beta\) be the real obtained by adding \(\alpha\) with \(\lambda\). Let \(B\) be set of digits obtained by \(r\)-adic representation of \(\beta\). Consequently, if we take alphabet \(A\) as input alphabet, then \(B\) is output alphabet. Also, according to [5], there is a transformation — invariant measure in \(B\) and since
\[
\nu_f \equiv \nu_\beta = \nu
\]

Hence by preceding theorem, normality of \(\nu\) implies normality of \(\alpha\).

ii) Since every digit of \(\beta\) occurs in a member of \([A]\), So probability occurrence of that digit in \(\beta\) is same as probability occurrence of that digit in member of \([A]\). The same argument holds in case of block of \(K\)-digits. Thus, \(\beta\) is normal to scale \(r\).

**Corollary** Let \(\beta = (\alpha_1, \alpha_2, \ldots, \alpha_k)\) be a normal \(K\)-tuple to scale \(r\).
1) every digit of a (1 e letter of input alphabet A) gives a unique —
digit of e (1 e letter of out put alphabet B)

2) Since the probability $\pi_i$ that digits (events of B) Will be a digit
(or block of digits) in r-adic scale of $\delta$ depends on whole set of digits of a
1: events of $A$. So we get a noise-less channel $[A, \pi, B]$ which is stationary.

Now according to Khinchin (1957), connecting a stationary source
with a stationary channel where alphabet of source coincides with input
ACKNOWLEDGEMENT

The author expresses his earnest gratitude to Dr M Dutta of Centre of Advanced Study in Applied Mathematics, Calcutta University for his kind help and guidance in preparation of this paper.

Received 3 VIII 1971

REFERENCES

2) KHINCHIN A I (1957) Mathematical foundations of Information Theory (Dover Publ)
3) SHANNON C E (1948) Mathematical Theory of Communications Bell Syst Tech J, 27
4) LUPS KEAH, MAHIELD (1969) Further examples of Non-normal numbers Publ Math Debrecen 16
5) COLENS, J (1961) Ziffern Versetzung in Q-adschiers Reaschers Math Zet, 78
6) S K MUKHERJEE: On applications of Information — theoretic notions in normal K-tuples Communicated