The Annals of Mathematical Statistics
1972, Vol. 43, No. 2, 541-552

THE MAXIMIZATION OF ENTROPY OF DISCRETE DENUMERABLY-VALUED RANDOM VARIABLES WITH KNOWN MEAN[1]

BY ROMAN A. MUREIKA

University of Alberta

1. Introduction. Let $S = \{u_k\}$ be a set of real numbers. A discrete random variable will be said to be S-valued if S represents the totality of all possible values of the random variable. If $P[X = u_k] = p_k$, then the entropy of X is defined to be $H_x = -\sum_k p_k \log p_k$ where the logarithm is taken to the base 2. It is of some interest to find the maximum value of H_x over the set of all S-valued random variables.

When S is a finite set consisting of n elements, $\max H_x = \log n$ [1]. However, if the set S is countably infinite, then an S-valued random variable may have infinite entropy. Since this is the case, it is natural to place some restrictions on the set of random variables and to then determine the maximum entropy. Again, if S is a finite set, this has been done and the result will be stated in Theorem 1.

Let $S = \{u_1, \cdots, u_n\}$ be a set of real numbers. Let $f_1, f_2, \cdots, f_m$ $(m < n)$ be m linearly independent real-valued functions. We define

$$Z(x_1, \cdots, x_m) = \sum_{k=1}^{n} \exp\{-\sum_{j=1}^{m} x_j f_j(u_k)\}$$

and $H = \max H_x$ where the maximum is taken over the set of S-valued random variables under the condition that $Ef_j = f_j^{(0)}$, $j = 1, 2, \cdots, m$, for a fixed collection of numbers $\{f_j^{(0)}\}$.

THEOREM 1. (i) $H = \sum_{j=1}^{m} \hat{x}_j f_j^{(0)} + \log Z(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_m)$ *where* $(\hat{x}_1, \cdots, \hat{x}_m)$ *is the unique solution of the set of equations*

$$\frac{\partial}{\partial x_j}[\log Z(x_1, \cdots, x_m)] = -f_j^{(0)}, \qquad j = 1, 2, \cdots, m$$

(ii) $H = -\sum_{k=1}^{n} p_k \log p_k$ *and* $\sum_{k=1}^{m} p_k f_j(u_k) = f_j^{(0)}$, $j = 1, 2, \cdots, m$ *if and only if*

$$p_k = \exp\{-\lambda - \sum_{j=1}^{m} \hat{x}_j f_j(u_k)\}, \qquad k = 1, 2, \cdots, n$$

where $\lambda = \log Z(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_m)$.

A proof of this result may be found in [2] or [3]. We will use it in the case $m = 1$ and $f(x) = x$.

Received January 30, 1970.

[1] This forms a part of the author's dissertation directed by Professor Eugene Lukacs and submitted in partial satisfaction of the requirements of a graduate program in Mathematics for the degree of Doctor of Philosophy at the Catholic University of America, Washington, D.C. This research was supported in part by NSF Grant GP-6175.

2. Relation of maximum entropy between finite and infinite cases. Let $S = \{u_1, u_2, \cdots, u_n, \cdots\}$. We will assume that S is an increasing sequence of points. H_∞ is defined as follows

$$H_\infty = \sup\{-\textstyle\sum_{k=1}^\infty p_k \log p_k \mid p_k \geqq 0,\ \sum_{k=1}^\infty p_k = 1,\ \sum_{k=1}^\infty p_k u_k = \bar{u}\}\ .$$

Then H_∞ is the supremum of entropy of all S-valued random variables with mean $\bar{u}$. We note here that H_∞ may be infinite. $\bar{u}$ will be assumed to be greater than u_1 since no S-valued random variable may have mean smaller than u_1 and $\bar{u} = u_1$ is trivial. The following theorem relates H_∞ to the maximum entropy of random variables taking values in finite subsets of S and having $\bar{u}$ as their mean value. Let $H_{S'}$ denote the maximum entropy of S'-valued random variables with mean $\bar{u}$.

THEOREM 2. *$H_\infty = \sup_{S' \subseteqq S} H_{S'}$ where S' is a finite subset of S.*

PROOF. Let $S' \subseteqq S$ be arbitrary. Since each S'-valued random variable is also S-valued, we see immediately that $H_{S'} \leqq H_\infty$. Hence, it follows that

$$\sup_{S' \subseteqq S} H_{S'} \leqq H_\infty\ . \tag{2.1}$$

We set $H = \sup H_{S'}$. It will be shown that equality must hold in (2.1) by assuming that $H < H_\infty$ and having this lead to a contradiction of the definition of H.

Let M be a real number such that $H < M < H_\infty$. From the definition of H_∞, there exists a set of probabilities $\{p_k\}$, $k = 1, 2, \cdots$ such that

$$-\textstyle\sum_{k=1}^\infty p_k \log p_k > M$$

and there exists an N such that $u_{N+1} > 0$ and

$$-\textstyle\sum_{k=1}^N p_k \log p_k > H\ .$$

Let

$$d_N = \bar{u} - \textstyle\sum_{k=1}^N p_k u_k > 0$$

and

$$q_N = 1 - \textstyle\sum_{k=1}^N p_k > 0\ .$$

To complete the proof, we need only find $q_1 \geqq 0$, $q_2 \geqq 0$ and $u_{n_1}, u_{n_2} \in \{u_n \mid n \geqq N+1\}$ such that

$$q_1 u_{n_2} + q_2 u_{n_2} = d_N \tag{2.2}$$

and

$$q_1 + q_2 = q_N \tag{2.3}$$

since we would then have a probability distribution for an S'-valued random variable with mean $\bar{u}$ where $S' = \{u_1, u_2, \cdots, u_N, u_{n_1}, u_{n_2}\} \subseteqq S$ and such that

$$H_{S'} \geqq -\textstyle\sum_{k=1}^N p_k \log p_k - q_1 \log q_1 - q_2 \log q_2 > H$$

which would be a contradiction of the definition of H since $H \geqq H_{S'}$.

Equations (2.2) and (2.3) have the solution

$$q_1 = (d_N - u_{n_2}q_N)/(u_{n_1} - u_{n_2})$$
$$q_2 = (u_{n_1}q_N - d_N)/(u_{n_1} - u_{n_2}) .$$

By choosing $u_{n_2} = u_{N+1}$ and $u_{n_1} \in \{u_n \mid n \geqq N+2\}$ where $u_{n_1} > d_N/q_n$, we see that $q_1 > 0$ and $q_2 > 0$.

Instead of Theorem 2, we will use the following corollary in our application of this result:

COROLLARY 1. *$H_\infty = \lim_{m\to\infty} H_m$ where H_m denotes the maximum entropy of a random variable taking values in the set $S_m = \{u_1, \cdots, u_m\}$ and having mean value $\bar{u}$.*

The proof follows trivially from Theorem 2.

LEMMA 1. *H_n and H_∞ are invariant under additive shift.*

PROOF. Here we need only note that if $\sum p_k u_k = \bar{u}$ then $\sum p_k(u_k + c) = \bar{u} + c$. Since the entropy is the same in each case, the supremum will give the same result.

3. Properties of $\hat{x}_n$. Let S be defined as in Section 2, but with the added assumption that $u_1 > 0$ and $u_n \to +\infty$ as $n \to +\infty$. This can be done without loss of generality by Lemma 1. Then, there exists an N such that $u_N \leqq \bar{u} < u_{N+1}$. To find H_∞, we can, by virtue of Corollary 1, examine the sequence H_n for $n \geqq N+1$. Let

$$Z_n(x) = \textstyle\sum_{k=1}^n e^{-u_k x}$$

and

$$\lambda_n(x) = \log Z_n(x) .$$

From Theorem 1, we see that there exists a unique $\hat{x}_n$ (for each $n \geqq N+1$) such that

$$H_n = \hat{x}_n \bar{u} + \lambda_n(\hat{x}_n) \tag{3.1}$$

which is the entropy of the random variable X where

$$P[X = u_k] = \exp\{-u_k \hat{x}_n - \lambda_n(\hat{x}_n)\} , \qquad k = 1, 2, \cdots, n$$

and $\hat{x}_n$ is the solution of the equation

$$Z_n'(x) + \bar{u} Z_n(x) = 0 . \tag{3.2}$$

In this way, we form a sequence $\{\hat{x}_n\}$, $n = N+1, N+2, \cdots$ where each $\hat{x}_n$ satisfies (3.2). A number of lemmas concerning $\{\hat{x}_n\}$ will be needed.

LEMMA 2. *There exists N_0 such that $\hat{x}_n > 0$ for all $n \geqq N_0$.*

PROOF. We can rewrite (3.2) for each n as,

$$-\textstyle\sum_{k=1}^n u_k e^{-u_k \hat{x}_n} + \bar{u} \sum_{k=1}^n e^{-u_k \hat{x}_n} = 0$$

or

$$L_n(\hat{x}_n) = R_n(\hat{x}_n) \tag{3.3}$$

where

$$L_n(x) = \sum_{k=1}^{N} (\bar{u} - u_k)e^{-u_k x}$$

and

$$R_n(x) = \sum_{k=N+1}^{n} (u_k - \bar{u})e^{-u_k x} .$$

Suppose that Lemma 2 is not true and that for each N_0, there exists $n_0 > N_0$ such that $\hat{x}_{n_0} \leqq 0$. Let $N_0 > N(u_{N+1} - u_1)/(u_{N+1} - \bar{u}) + N$. Then, by our assumption, there exists $n_0 \geqq N_0$ such that $\hat{x}_n \leqq 0$ and

$$L_{n_0}(\hat{x}_{n_0}) \leqq N(\bar{u} - u_1) \exp[-\hat{x}_{n_0} u_N]$$

and

$$R_{n_0}(\hat{x}_{n_0}) \geqq (n_0 - N)(u_{N+1} - \bar{u}) \exp[-\hat{x}_{n_0} u_{N+1}] .$$

From (3.3), we see that

$$N(\bar{u} - u_1) \geqq (n_0 - N)(u_{N+1} - \bar{u}) \exp[-\hat{x}_{n_0}(u_{N+1} - u_N)] .$$

Since $\hat{x}_{n_0} \leqq 0$,

$$N(\bar{u} - u_1) \geqq (n_0 - N)(u_{N+1} - \bar{u})$$

or

$$n_0 \leqq N(\bar{u} - u_1)/(u_{N+1} - \bar{u}) + N < N_0$$

which contradicts the fact that $n_0 \geqq N_0$ so that $\hat{x}_n > 0$ for n sufficiently large.

LEMMA 3. *There exists an $\varepsilon > 0$ and an N_1 such that for all $n \geqq N_1$, $\hat{x}_n \geqq \varepsilon > 0$.*

PROOF. It is easy to see that $L_n(\hat{x}_n)$ and $R_n(\hat{x}_n)$ are bounded sequences of numbers, since

$$\begin{aligned} L_n(\hat{x}_n) &= \sum_{k=1}^{N} (\bar{u} - u_k)e^{-u_k \hat{x}_n} \\ &\leqq \sum_{k=1}^{N} (\bar{u} - u_k) = C' \end{aligned}$$

for $n \geqq N_0$ where N_0 is given by Lemma 2. Hence,

$$0 \leqq L_n(\hat{x}_n) \leqq C = \max\{L_{N+1}(\hat{x}_{N+1}), \cdots, L_{N_0}(\hat{x}_{N_0}), C'\} \tag{3.4}$$

and it follows from (3.3) that

$$0 \leqq R_n(\hat{x}_n) \leqq C . \tag{3.5}$$

Thus, from (3.4), (3.5) and the definitions of $L_n(\hat{x}_n)$ and $R_n(\hat{x}_n)$, it follows that for each n,

$$C \geqq (\bar{u} - u_j)e^{-u_j \hat{x}_n} , \qquad 1 \leqq j \leqq N$$

and

$$C \geqq (u_j - \bar{u})e^{-u_j \hat{x}_n} , \qquad N + 1 \leqq j \leqq n$$

or

$$C \geqq |u_j - \bar{u}|e^{-u_j \hat{x}_n} , \qquad 1 \leqq j \leqq n .$$

Thus,

$$\hat{x}_n \geqq \frac{1}{u_j} \log \frac{|u_j - \bar{u}|}{C}, \qquad 1 \leqq j \leqq n. \tag{3.6}$$

Since $u_j \to +\infty$, there exists an N_1 such that $(u_{N_1} - \bar{u})/C > 1$ so that $\log |u_{N_1} - \bar{u}|/C > 0$.

Let
$$\varepsilon = u_{N_1}^{-1} \log |u_{N_1} - \bar{u}|/C > 0 .$$
Since (3.6) holds for each $j \leqq n$, fixing $j = N_1$, implies that (3.6) holds for all $n \geqq N_1$. Hence $\hat{x}_n \geqq \varepsilon > 0$ for all $n \geqq N_1$.

We will need Lemma 3 in the proof of Theorem 3 in the next section.

The following result illustrates some of the relationship between $\{\hat{x}_n\}$ and H_∞.

LEMMA 4. *If* $\limsup_n \hat{x}_n = +\infty$ *then* $H_\infty = +\infty$.

PROOF. From (2.1), we see that

$$H_n = \hat{x}_n \bar{u} + \lambda_n(\hat{x}_n) = \log \textstyle\sum_{k=1}^{N} \exp\{(\bar{u} - u_k)\hat{x}_n\} .$$

Since each of the terms in the above sum is positive, it follows that

$$H_n \geqq \log \exp[(\bar{u} - u_1)\hat{x}_n] = (\bar{u} - u_1)\hat{x}_n$$

so that $H_\infty \geqq (\bar{u} - u_1)\hat{x}_n$ for each $n \geqq N + 1$.

But $\limsup_n \hat{x}_n = +\infty$ implies immediately that $H_\infty = +\infty$.

4. Unbounded sets. We may now proceed to find necessary and sufficient conditions on S for H_∞ to be finite. It will be seen that the important factor for this will be the rate at which the sequence $\{u_n\}$ converges infinity. Theorem 3 gives a sufficient condition for H_∞ to be infinite.

THEOREM 3. *If* $\limsup_n \log n/u_n = +\infty$, *then* $H_\infty = +\infty$.

PROOF. From (3.5), we note that there exists a $C > 0$ such that for every $n \geqq N + 1$, $C \geqq R_n(\hat{x}_n)$ or

$$Ce^{\bar{u}\hat{x}_n} \geqq \textstyle\sum_{k=N+1}^{n} u_k' e^{-\hat{x}_n u_k'} \tag{4.1}$$

where $u_k' = u_k - \bar{u}$, $k = N + 1, N + 2, \cdots$. Furthermore, from Lemma 3, there exists an N_1 and $\varepsilon > 0$ such that $\hat{x}_n > \varepsilon$ or $\varepsilon^{-1} > \hat{x}_n^{-1}$ for $n \geqq N_1$. Since u_k' is increasing, there exists an integer M such that

$$u_k' \geqq \varepsilon^{-1} > \hat{x}_n^{-1}, \qquad k \geqq M, \quad n \geqq N_1 . \tag{4.2}$$

We note that ue^{-uk} is a monotonically decreasing function in u for $u > k^{-1}$ so that for n fixed but arbitrary $u_k' e^{-\hat{x}_n u_k'}$ is strictly decreasing in k for $u_k' > \hat{x}_n^{-1}$. (4.1) and (4.2) then imply that for $n \geqq A = \max\{N_1, M\}$,

$$\begin{aligned} Ce^{\bar{u}\hat{x}_n} &\geqq \textstyle\sum_{k=A}^{n} u_k' e^{-u_k'\hat{x}_n} \geqq (n - A + 1) u_n' e^{-\hat{x}_n u_n'} \\ &= (n - A + 1)(u_n - \bar{u}) \exp[-\hat{x}_n(u_n - \bar{u})] \end{aligned}$$

or

$$e^{\hat{x}_n u_n} \geqq (n - A + 1)(u_n - \bar{u})/C \qquad n \geqq A .$$

Hence

$$\hat{x}_n \geqq \frac{\log n}{u_n} + \frac{1}{u_n} \log \frac{(n - A + 1)(u_n - \bar{u})}{nC}$$
$$= \frac{\log n}{u_n} + o(1) \qquad \text{as} \quad n \to +\infty .$$

This means that $\limsup_n \hat{x}_n = +\infty$ and Lemma 4 immediately implies that $H_\infty = +\infty$.

We are now ready to state the following result which together with Theorem 3 provides the necessary and sufficient conditions for H_∞ to be finite.

THEOREM 4. *Let* $\limsup_n \log n/u_n < +\infty$. *Then* $H_\infty < +\infty$.

PROOF. The first step in the proof is to show that the sequence $\{\hat{x}_n\}$ is bounded. We assume that this is not true and show that this contradicts the inequality $\bar{u} > u_1$.

From Lemma 2, it is seen that for n sufficiently large, $\hat{x}_n > 0$. Hence, $\{\hat{x}_n\}$ unbounded implies that there exists a subsequence $\{\hat{x}_{n_m}\}$ such that $\hat{x}_{n_m} \to +\infty$ as $j \to +\infty$. We see from Theorem 1, that H_{n_j} is achieved by the set of probabilities

$$(4.3) \qquad P_{l,n_j} = \exp\{-u_l \hat{x}_{n_j} - \lambda_{n_j}(\hat{x}_{n_j})\}$$
$$= \exp[-u_l \hat{x}_{n_j}] / \textstyle\sum_{k=1}^{n_j} \exp[-u_k \hat{x}_{n_j}] , \qquad l = 1, 2, \cdots, n_j$$

where $\sum_{l=1}^{n_j} P_{l,n_j} u_l = \bar{u}$.

Then, for $l > 1$,

$$(4.4) \qquad P_{l,n_j} \leqq \exp[-u_l \hat{x}_{n_j}] / \exp[-u_1 \hat{x}_{n_j}]$$
$$= \exp\{(u_1 - u_l)\hat{x}_{n_j}\} \to 0 \qquad \text{as} \quad n_j \to +\infty$$

since $\hat{x}_{n_j} \to +\infty$.

Let $0 < \varepsilon < 1$ be an arbitrary real number. Then there exists an N' such that for $n_j > N'$, $p_{2,n_j} < \varepsilon$. This implies that

$$\exp[-u_2 \hat{x}_{n_j}] < \varepsilon \textstyle\sum_{k=1}^{n_j} \exp[-u_k \hat{x}_{n_j}] \qquad n_j > N'$$

so that for $l \geqq 2$,

$$(4.5) \qquad \exp[-u_l \hat{x}_{n_j}] = (\exp[-u_2 \hat{x}_{n_j}])^{u_l/u_2}$$
$$< \varepsilon^{u_l/u_2} \{\textstyle\sum_{k=1}^{n_j} \exp[-u_k \hat{x}_{n_j}]\}^{u_l/u_2} \qquad n_j > N' .$$

From (4.3), (4.4), (4.5) and the fact that $p_{1,n_j} \leqq 1$, for all n; it is seen that

$$(4.6) \qquad \bar{u} < u_1 + \textstyle\sum_{k=2}^{n_j} \varepsilon^{u_k/u_2} u_k \{\sum_{l=1}^{n_j} \exp[-u_l \hat{x}_{n_j}]\}^{(u_k/u_2 - 1)} .$$

Now, $\limsup_n \log n/u_n < +\infty$ implies that there exists a $c > 0$ such that for all n,

$$\log n/u_n \leqq c . \tag{4.7}$$

This in turn implies that $e^{-cu_n} \leqq n^{-1}$ or $e^{-2cu_n} \leqq n^{-2}$ so that

$$\sum_{n=1}^{\infty} e^{-2cu_n} \leqq \sum_{n=1}^{\infty} n^{-2} < +\infty .$$

Since $\hat{x}_{n_j} \to +\infty$, there exists an N'' such that $\hat{x}_{n_j} > 2c$ for $n_j > N''$. Hence for $n_j > N''$

$$\sum_{l=1}^{n_j} \exp[-u_l \hat{x}_{n_j}] \leqq \sum_{l=1}^{\infty} \exp[-u_l \hat{x}_{n_j}] \leqq \sum_{l=1}^{\infty} e^{-2cu_l} = \beta < +\infty . \tag{4.8}$$

We let $\eta = -\log \varepsilon > 0$ and $\mu = \log \beta$. Then using (4.8) we see that (4.6) becomes

$$\begin{aligned} \bar{u} &< u_1 + \sum_{k=2}^{n_j} u_k \exp\left\{-\eta \frac{u_k}{u_2} + \left(\frac{u_k}{u_2} - 1\right)\mu\right\} \\ &= u_1 + e^{-\mu} \sum_{k=2}^{n_j} u_k \exp\{-u_k(\eta - \mu)/u_2\} , \quad n_j \geqq M = \max\{N', N''\} . \end{aligned} \tag{4.9}$$

Let $\tau(\varepsilon) = (\eta - \mu)/u_2 = -\log \varepsilon\beta/u_2$. Since β is independent of ε, we see that $\tau(\varepsilon) \to +\infty$ as $\varepsilon \to 0$. We limit the range of ε to $0 < \varepsilon < \varepsilon'$ where ε' is such that $\tau(\varepsilon') > C/\log 3$ where C is determined from (4.7). By noting that $u_k \exp[-\tau(\varepsilon)u_k]$ is decreasing in k for $u_k > \tau(\varepsilon')^{-1}$, we see from (4.7) and (4.9) that for $n_j \geqq M$,

$$\begin{aligned} \bar{u} &< u_1 + e^{-\mu}\left[u_2 \exp[-u_2\tau(\varepsilon)] + \sum_{k=3}^{n_j} \frac{\log k}{C} \exp\left\{-\left[\frac{\log k}{C}\right]\tau(\varepsilon)\right\}\right] \\ &\leqq u_1 + e^{-\mu}\left[u_2 \exp[-u_2\tau(\varepsilon)] + \frac{1}{C} \sum_{k=3}^{n_j}\left(\frac{1}{k}\right)^{\tau(\varepsilon)/C-1}\right], \qquad \varepsilon \leqq \varepsilon' , \quad n_j \geqq M \end{aligned}$$

since $\log k \leqq k$ for $k > 2$.

Hence, if ε is chosen sufficiently small so that $\tau(\varepsilon) > 3C$, we see that

$$\begin{aligned} \bar{u} &< u_1 + e^{-\mu}[u_2 \exp[-u_2\tau(\varepsilon)] + C^{-1} \sum_{k=3}^{n_j} (k^{-1})^{\tau(\varepsilon)/C-1}] \\ &\leqq u_1 + e^{-\mu}[u_2 \exp[-u_2\tau(\varepsilon)] + C^{-1} \sum_{k=3}^{\infty} (k^{-1})^{\tau(\varepsilon)/C-1}] \\ &= u_1 + g(\varepsilon) < +\infty \qquad\qquad n_j \geqq M . \end{aligned}$$

Since $g(\varepsilon)$ converges for at least one value of ε and since $\tau(\varepsilon) \to +\infty$ as $\varepsilon \to 0$, we see that $g(\varepsilon) \to 0$ as $\varepsilon \to 0$ and it follows immediately that $\bar{u} \leqq u_1$. This contradicts the fact that $\bar{u} > u_1$ so that $\{\hat{x}_n\}$ is bounded. For each n, we note that

$$\sum_{k=1}^{N} (\bar{u} - u_k)e^{-u_k\hat{x}_n} = \sum_{u=N+1}^{n} (u_k - \bar{u})e^{-u_k\hat{x}_n} . \tag{4.10}$$

Since the sequence $\{\hat{x}_n\}$ is bounded we see that for some constant C and each $n \geqq N + 1$,

$$C \geqq \sum_{k=1}^{N} (\bar{u} - u_k)e^{-u_k\hat{x}_n} = \sum_{k=N+1}^{n} (u_k - \bar{u})e^{-u_k\hat{x}_n} \geqq (u_{N+1} - \bar{u}) \sum_{k=N+1}^{n} e^{-u_k\hat{x}_n}$$

or

$$C/(u_{N+1} - \bar{u}) \geqq \sum_{k=N+1}^{n} e^{-u_k \hat{x}_n} .$$

Hence for some constant C_1,

$$C_1 \geqq \sum_{k=1}^{n} e^{-u_k \hat{x}_n} \qquad n \geqq N + 1 . \tag{4.11}$$

But this implies that

$$H_n = \hat{x}_n \bar{u} + \log \sum_{k=1}^{n} e^{-u_k \hat{x}_n} \leqq \hat{x}_n \bar{u} + \log C_1 \leqq M\bar{u} + \log C_1 < +\infty$$

where M is an upper bound for the sequence $\{\hat{x}_n\}$.
From Corollary 1, we see that $H_\infty = \lim_{n\to\infty} H_n < +\infty$, which proves the theorem.

The previous theorem gives a sufficient condition for H_∞ to be finite but does not provide means to find this value apart from evaluating the limit. Theorem 5 will provide the means to do this under certain conditions. We will first need the following lemma:

LEMMA 5. *Let* $\limsup_n \log n/u_n = \gamma (\leqq +\infty)$. *Then* $\sum e^{-u_k x}$ *and* $\sum u_k{}^n e^{-u_k x}$ *diverge for* $x < \gamma$ *and converge for* $x > \gamma$ *for each integer* n.

PROOF. We first prove that $\sum_{k=1}^{\infty} e^{-u_k x}$ converges for $x > \gamma$. Let $\varepsilon > 0$ be arbitrary. Then, by the definition of γ, there exists an $n_0(\varepsilon)$ such that

$$\log k/u_k \leqq \gamma + \varepsilon , \qquad k \geqq n_0(\varepsilon) .$$

Hence, for any $x > 0$,

$$\frac{x}{\gamma + \varepsilon} \log k \leqq u_k x , \qquad k \geqq n_0(\varepsilon)$$

or

$$e^{-u_k x} \leqq (k^{-1})^{x/\gamma+\varepsilon} , \qquad k \geqq n_0(\varepsilon) .$$

Thus, for each $m > n_0(\varepsilon)$

$$\sum_{k=n_0}^{m} e^{-u_k x} \leqq \sum_{k=n_0}^{m} (k^{-1})^{x/\gamma+\varepsilon} < \sum_{k=n_0}^{\infty} (k^{-1})^{x/\gamma+\varepsilon} .$$

This implies that $\sum_{k=n_0}^{\infty} e^{-u_k x} < +\infty$ for $x > \gamma + \varepsilon$. Since ε is arbitrary, we see that the series converges for each $x > \gamma$. To show that $\sum_{k=1}^{\infty} e^{-u_k x}$ diverges for $x < \gamma$ we will need the following lemma:

LEMMA (Kronecker's lemma). *Let* $\{a_k\}$ *be a sequence of non-decreasing real numbers such that* $\lim_{k\to+\infty} a_k = +\infty$. *Let* x_k *be an arbitrary sequence of real numbers. If* $\sum_{k=1}^{\infty} x_k/a_k$ *converges, then* $\lim_{n\to\infty} \sum_{k=1}^{n} x_k/a_n = 0$.

We can assume without loss of generality that $\gamma > 0$ since for $x < 0$, $\sum_{k=1}^{\infty} e^{-u_k x}$ diverges. Suppose for purpose of reaching a contradiction that $\sum_{k=1}^{\infty} e^{-u_k x_0} < +\infty$ for some x_0 such that $0 < x_0 < \gamma$. Let $a_k = e^{u_k x_0}$ and $x_k = 1$. Then

$$\sum_{k=1}^{\infty} x_k/a_k = \sum_{k=1}^{\infty} e^{-u_k x_0} < +\infty .$$

By Kronecker's lemma this implies that

$$\lim_{n\to\infty} a_k^{-1} \sum_{k=1}^{n} x_k = 0$$

or

$$\lim_{n\to\infty} e^{-u_k x_0} n = 0 .$$

Hence for $n > N_0$, $e^{-u_k x_0} \cdot n < 1$.

Taking the logarithm of both sides, we see that this implies that

$$\log n/u_n < x_0 < \gamma \qquad n > N_0$$

which contradicts the definition of γ. It can be noted here that nothing can be said about convergence or divergence of $\sum_{k=1}^{\infty} e^{-u_k \gamma}$ since examples of both cases exist. The rest of the lemma now follows easily from what has been shown.

$$\sum_{k=1}^{\infty} u_k^n e^{-u_k x} > u_1^n \sum_{k=1}^{\infty} e^{-u_k x}$$

implies that the left-hand side diverges for $x < \gamma$. To prove convergence for $x > \gamma$, we let x' be some number such that $\gamma < x' < x$. Then

$$\sum_{k=1}^{\infty} e^{-u_k x'} = \sum_{k=1}^{\infty} e^{-u_k x} \exp[u_k(x - x')] \geqq \sum_{k=1}^{\infty} \frac{u_k^n (x - x')^n}{n!} e^{-u_k x}$$

$$= \frac{(x - x')^n}{n!} \sum_{k=1}^{\infty} u_k^n e^{-u_k x}$$

since $e^x \geqq x^n/n!$ for $x > 0$. The left-hand side of the above inequality converges for $x' > \gamma$ and hence so does the right-hand side.

THEOREM 5. *Suppose that* $\limsup_n \log n/u_n = \gamma < +\infty$ *and* $\sum_{k=1}^{\infty} u_k e^{-u_k \gamma} = +\infty$. *Then there exists an* $x_\infty < +\infty$ *such that*

(i) $x_\infty = \lim_{n\to\infty} \hat{x}_n$

(ii) $H_\infty = \bar{u} x_\infty + \lambda(x_\infty)$ *where*

$$\lambda(x_\infty) = \log \sum_{k=1}^{\infty} e^{-u_k x_\infty} \qquad \textit{and}$$

(iii) H_∞ *is attained by the random variable* X *where* $P[X = u_k] = \exp\{-u_k x_\infty - \lambda(x_\infty)\}$ *for* $k = 1, 2, \cdots$.

PROOF. In the proof of Theorem 4, we have already seen that the sequence $\{\hat{x}_n\}$ is bounded. Now every bounded sequence of real of numbers has an accumulation point. Let x_0 be such a point. We will show first that $x_0 > \gamma$. Let $\{\hat{x}_{n_j}\}$ be a subsequence of $\{\hat{x}_n\}$ which converges to x_0. Then for arbitrary $\varepsilon > 0$, there exists an $n(\varepsilon)$ such that

$$x_{n_j} \leqq x_0 + \varepsilon \qquad n_j \geqq n_0(\varepsilon) .$$

Combining statements (4.10) and (4.11) we see that there exists a constant K such that for each n,

$$K \geqq \sum_{k=1}^{n} u_k e^{-u_k \hat{x}_n}$$

so that for $n_j \geqq n_0(\varepsilon)$

$$K \geqq \sum_{k=1}^{n_j} u_k \exp[-u_k(x_0 + \varepsilon)]$$

for each $\varepsilon > 0$. Thus letting $n_j \to +\infty$ gives

$$K \geqq \sum_{k=1}^{\infty} u_k \exp[-u_k(x_0 + \varepsilon)] , \qquad \varepsilon > 0 .$$

Now, for each fixed n

$$K \geqq \sum_{k=1}^{n} u_k \exp[-u_k(x_0 + \varepsilon)] \geqq e^{-u_n \varepsilon} \sum_{k=1}^{n} u_k e^{-u_k x_0} , \qquad \varepsilon > 0$$

or

$$e^{u_n \varepsilon} K \geqq \sum_{k=1}^{n} u_k e^{-u_k x_0} , \qquad \varepsilon > 0 .$$

Letting $\varepsilon \to 0$ gives

$$K \geqq \sum_{k=1}^{n} u_k e^{-u_k x_0} .$$

This is true for each n, so we see that

$$\sum_{k=1}^{\infty} u_k e^{-u_k x_0} < +\infty .$$

Since we have assumed that $\sum_{k=1}^{\infty} u_k e^{-u_k \gamma} = +\infty$ it is easily seen that $x_0 > \gamma$. Next we show that

$$\sum_{k=N+1}^{n_j} (u_k - \bar{u}) \exp[-u_k \hat{x}_{n_j}] \to \sum_{k=N+1}^{\infty} (u_k - \bar{u}) e^{-u_k x_0}$$

where $\hat{x}_{n_j} \to x_0$.

Let $\varepsilon > 0$ be given. Then there exists N_1 such that for all $n \geqq N_1$,

$$\sum_{k=n}^{\infty} (u_k - \bar{u}) \exp[-u_k(x_0 + \gamma)/2] < \varepsilon/3$$

since the series converges for all $x > \gamma$. We note that this also implies that $\sum_{k=n}^{\infty} (u_k - \bar{u}) e^{-u_k x} < \varepsilon/3$ for all $x \geqq (x_0 + \gamma)/2$. Thus for $n \geqq N_1$

$$\begin{aligned} &\left|\sum_{k=N+1}^{n_j} (u_k - \bar{u}) \exp[-u_k \hat{x}_{n_j}] - \sum_{k=N+1}^{\infty} (u_k - \bar{u}) e^{-u_k x_0}\right| \\ &\quad \leqq \left|\sum_{k=N+1}^{N_1} (u_k - \bar{u})(\exp[-u_k \hat{x}_{n_j}] - e^{-u_k x_0})\right| \\ &\qquad + \left|\sum_{k=N_1+1}^{n_j} (u_k - \bar{u}) \exp[-u_k \hat{x}_{n_j}]\right| + \left|\sum_{k=N_1+1}^{\infty} (u_k - \bar{u}) e^{-u_k x_0}\right| . \end{aligned}$$

Since $\hat{x}_{n_j} \to x_0$, there exists an N_2 such that for $x_{n_j} \geqq N_2$, $\hat{x}_{n_j} \geqq (x_0 + \gamma)/2$. Similarly there exists an N_3 such that the first term on the right-hand side above is smaller than $\varepsilon/3$ for $n_j \geqq N_3$. Hence for $n_j \geqq \max(N_1, N_2, N_3)$ we see that the right-hand side may be made less than ε. It can now be shown that the limit point is unique.

Now, every bounded sequence of real numbers has an accumulation point. Let x_1 be such a point. We will show that it is unique: Suppose $x_2 \neq x_1$ is also an accumulation point. We may assume without loss of generality that $x_2 > x_1$. There must exist subsequences $\{\hat{x}_{n_j}\}$ and $\{\hat{x}_{m_j}\}$ of $\{\hat{x}_n\}$ which converge to x_1 and x_2, respectively. Since

$$\sum_{k=1}^{n} (\bar{u} - u_k) e^{-u_k \hat{x}_n} = \sum_{k=N+1}^{n} (u - \bar{u}) e^{-u_k \hat{x}_n} \tag{4.12}$$

for each n, this is true in particular each n_j and m_j. Taking the limits of both sides in (4.12) as $j \to +\infty$, we see that

$$\sum_{k=1}^{N} (\bar{u} - u_k)e^{-u_k x_1} = \sum_{k=N+1}^{\infty} (u_k - \bar{u})e^{-u_k x_1} \tag{4.13}$$

and

$$\sum_{k=1}^{N} (\bar{u} - u_k)e^{-u_k x_2} = \sum_{k=N+1}^{\infty} (u_k - \bar{u})e^{-u_k x_2} . \tag{4.14}$$

Multiplying (4.13) by $e^{\bar{u}x_1}$ and (4.14) by $e^{\bar{u}x_2}$, and taking the difference, we get

$$\begin{aligned} \sum_{k=1}^{N} (\bar{u} - u_k)[\exp[(\bar{u} - u_k)x_1] - \exp[(\bar{u} - u_k)x_2]] \\ = \sum_{k=N+1}^{\infty} (u_k - \bar{u})[\exp[(\bar{u} - u_k)x_1] - \exp[(\bar{u} - u_k)x_2]] . \end{aligned} \tag{4.15}$$

Since $x_2 > x_1$, we see that the left side of (4.15) is negative and the right side is positive. This is a contradiction and the accumulation point is unique. We now let $x_\infty = \lim_{n\to\infty} x_n$ and result (i) has been shown. Now,

$$H_n = \hat{x}_n \bar{u} + \log \sum_{k=1}^{n} e^{-u_k \hat{x}_n}$$

implies that, as $n \to \infty$

$$H_n \to x_\infty \bar{u} + \log \sum_{k=1}^{\infty} e^{-u_k x_\infty} = \bar{H} .$$

To see that $\bar{H}$ is finite, we note from (4.11) that

$$\sum_{k=N+1}^{\infty} e^{-u_k x_\infty} \leqq [u_{N+1} - \bar{u}]^{-1} \sum_{k=N+1}^{\infty} (u_k - \bar{u})e^{-u_k x_\infty} < +\infty .$$

We apply Corollary 1 and see that

$$H_\infty = \lim_{n\to\infty} H_n = \bar{H} < +\infty$$

which is result (ii). Finally, statement (iii) can be varified by direct calculation. We can now use Lemma 1 to extend the results of Theorems 3, 4, and 5:

THEOREM 6. *Let $S = \{u_n\}$, $n = 1, 2, \cdots$ be an arbitrary increasing sequence of real numbers such that $\lim_{n\to\infty} u_n = +\infty$ and let $u_1 < \bar{u}$. Then H_∞ is finite if and only if* $\limsup_n \log n/u_n < +\infty$.

The proof of this statement will follow directly from Lemma 1. We note also that it has been no restriction to consider only increasing sequences of real numbers. If $u_n \to -\infty$ as $n \to +\infty$, we need only consider the sequence of points $v_k = -u_k$ and $\bar{v} = -\bar{u}$ and apply the previous results to this new set of points.

5. Bounded sequences. If the set of points $\{u_k\}$ is bounded from above as well as below, the situation becomes somewhat simpler. Theorem 7 gives a complete solution to this problem.

THEOREM 7. *Let $S = \{u_n\}$ be a strictly increasing sequence of real numbers such that* $\lim_{k\to\infty} u_k = C < +\infty$. *Let $\bar{u}$ be such that $u_1 < \bar{u} < c$. Then $H_\infty = +\infty$.*

PROOF. We may assume without loss of generality that $u_1 > 0$. Since $u_k \to C$ as $k \to +\infty$, there exists an N such that $u_N \leqq \bar{u} < u_{N+1}$.

Let j be an arbitrary positive integer. We define $\bar{v}_j = (u_{N+1} + \cdots + u_{N+j})/j$ and $p = (\bar{u} - u_1)/j(\bar{v}_j - u_1)$.

Letting

$$\begin{aligned} p_k &= 1 - jp && k = 1 \\ &= p && N + 1 \leqq k \leqq N + j \\ &= 0 && \text{otherwise}\,, \end{aligned}$$

we see that

$$\textstyle\sum_{k=1}^{\infty} p_k = (1 - jp) + \sum_{k=1}^{j} p_{N+k} = 1 - jp + jp = 1$$

and

$$\begin{aligned} \textstyle\sum_{k=1}^{\infty} p_k u_k &= (1 - jp)u_1 + \textstyle\sum_{k=1}^{j} u_{N+k}\, p \\ &= (1 - jp)u_1 + pj\bar{v}_j \\ &= u_1 + jp(\bar{v}_j - u_1) = u_1 + \frac{\bar{u} - u_1}{\bar{v}_j - u_1}(\bar{v}_j - u_1) = \bar{u}\,. \end{aligned}$$

Since $0 \leqq p_k \leqq 1$ for all k, we see that

$$\begin{aligned} H_\infty &\geqq -\textstyle\sum_{k=1}^{\infty} p_k \log p_k \geqq -\sum_{k=1}^{j} p_{N+k} \log p_{N+k} \\ &= -jp \log p = \frac{\bar{u} - u_1}{\bar{v}_j - u_1}\left\{\log j + \log \frac{\bar{v}_j - u_1}{\bar{u} - u_1}\right\}. \end{aligned}$$

Since $\bar{u} < \bar{v}_j < c$ for all j and since j is arbitrary we see that $H_\infty = +\infty$. More general bounded sets S may be treated by choosing appropriate subsequences of S and applying the results of Theorem 7 to show that $H_\infty = +\infty$ for S.

Acknowledgment. The author wishes to thank the associate editor and the referee for their remarks and suggestions.

REFERENCES

[1] FEINSTEIN, A. (1958). *Foundations of Information Theory*. McGraw-Hill, New York.

[2] JAYNES, E. T. (1957). Information theory and statistical mechanics. *Phys. Rev.* **106** 620–630.

[3] TRIBUS, M. (1962). The use of the maximum entropy estimate in *Reliability, Recent Developments in Information and Decision Processes*, ed. by R. E. Machol and P. Gray. Macmillan, New York.