Boletín de Matemáticas
Vol. XIV , Nos. 1,2,3, (1980) pags, 68-106

UNA FUNCION CUASI-LOGARITMICA EN LA TEORIA DE LA INFORMACION

Jaime HERNANDEZ G.

O. RESUMEN

Si de un conjunto de N objetos se desea seleccionar uno en particular , este proceso se puede asociar con la obtención de una cantidad de información , la cual se puede calcular por medio de la conocida fórmula de Shannon . La fórmula de Shannon hace referencia al cambio ocurrido entre un estado inicial de incertidumbre y un estado final de conocimiento . En la práctica , sinembargo , algunos métodos son más efectivos que otros para proporcionar una determinada información . Se investiga en este artículo un método particular de selección : el binario , en el cual se evalúa el número de alternativas que en promedio se deben sortear para alcanzar un objeto del conjunto por dicho procedimiento . La función $H(N)$ que da el promedio de alternativas en términos del número N de

objetos se compara con la información logarítmica asociada con la misma selección, y se encuentra que ambas funciones coinciden cuando N es una potencia natural de 2. Se hallan algunas propiedades de $H(N)$, las cuales coinciden con propiedades fundamentales de la función logaritmo.

1. FORMULA DE SHANNON

Cuando se tiene un conjunto de N elementos de los cuales se selecciona uno, esta operación equivale a remover una incertidumbre, es decir, a adquirir información. Si los N elementos pueden ser escogidos con igual probabilidad, la información adquirida depende exclusivamente de N y es una función creciente de este número.

En la literatura abundan presentaciones y desarrollos más o menos detallados del cálculo de tal información (1), (2). Aquí nos limitaremos a dar un argumento heurístico para llegar a este bien conocido resultado en forma rápida.

Supongamos que cada uno de los elementos del conjunto ha sido numerado utilizando la notación binaria.

Escoger un miembro del conjunto equivale entonces a señalar el numeral que codifica al elemento en cuestión . Cada una de las cifras binarias de tal numeral sólo puede valer cero o uno , y la información adquirida se puede hacer igual a la suma de las informaciones adquiridas al seleccionar las cifras por separado .

Se llama un <u>bit</u> (<u>bi</u>nary digi<u>t</u>) la información asociada con la selección de uno entre dos objetos igualmente probables en cuanto a su escogencia . Si se tienen 2^M objetos , deberán seleccionarse independientemente M cifras del numeral binario que le corresponde al elegido , es decir , la información asociada con este proceso es M bits . Sea $I(2^M)$ la información asociada con la selección de uno entre estos 2^M objetos . Entonces podemos escribir :

$$I(2^M) = M . \qquad (1)$$

Esta fórmula también se puede escribir en la forma

$$I(2^M) = \log_2 2^M , \qquad (2)$$

y entonces es usual generalizarla (y se puede demostrar que es correcto hacerlo) para obtener

$$I(N) = \log_2 N , \qquad (3)$$

para cualquier N , no necesariamente una potencia natural de dos . Esta información estará medida en bits .

2. SELECCION BINARIA

La fórmula de Shannon expuesta en el parágrafo precedente permite el cálculo de la información asociada con la selección de un evento entre N que son igualmente probables , independientemente del proceso empleado para realizar la escogencia . En el presente aparte describiremos un método particularmente simple para hallar o escoger uno entre los N objetos dados .

El método es el siguiente : dividir el conjunto dado en dos partes que tengan igual número de objetos o difieran a lo sumo en una unidad , y decidir en cuál de los dos subconjuntos se encuentra el elemento buscado (1) ; fraccionar el conjunto favorecido en la misma forma , y así sucesivamente , hasta que quede un solo objeto que es el buscado .

---

(1) En este contexto no nos interesa el criterio para decidir en cuál conjunto se encuentra el elemento buscado .

Se puede lograr una descripción gráfica de este proceso si se piensa que es equivalente a separar el conjunto inicial en parejas (posiblemente quede un elemento aislado) y escoger un elemento de cada pareja para formar el subconjunto favorecido al cual se le hará el mismo fraccionamiento , y así hasta el final . Cuando queda un elemento aislado , éste siempre se incluirá en el subconjunto favorecido hasta que el número de elementos de dicho subconjunto sea par .

Por ejemplo , si tenemos 9 elementos (nivel cero) ,

Figura 1

dividimos el conjunto por parejas (más un elemento) y de cada pareja seleccionamos uno de los objetos (nivel 1) :

Figura 2

Ahora el conjunto favorecido consta de cinco elementos (no nos interesa saber cuáles), los cuales separamos por parejas y de cada una de éllas escogemos uno (nivel 2):

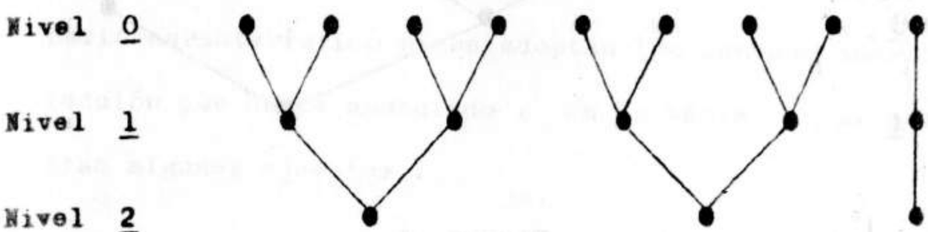


Figura 3

El nivel 2 tiene tres elementos de los cuales formamos parejas y escogemos nuevamente para obtener el nivel 3, y finalmente el nivel 4 con un solo elemento. Todo el proceso queda ilustrado en la figura 4.

Obsérvese que para mantener la simplicidad del fraccionamiento, siempre procedemos en el mismo orden para la formación de parejas (de izquierda a derecha o viceversa).

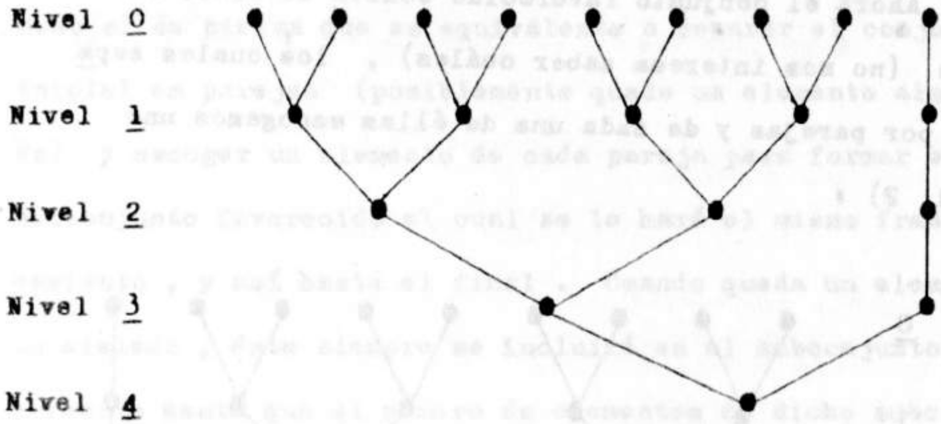


Figura 4

3. ARBOL BINARIO

Hemos obtenido que el proceso de selección binario se puede identificar con un árbol de cada uno de cuyos vértices (excepto del más bajo) se desprenden tres ramas : dos hacia el nivel anterior y una hacia el nivel posterior . Este tipo de árbol recibe el nombre de <u>árbol</u> <u>binario</u> . Observamos que cada vértice representa una decisión tomada entre dos alternativas y que , por lo tanto , cuando queda un elemento aislado , éste no representa una alternativa sino en la medida en que encontremos otro con el cual aparearlo . En el ejemplo , se-

gún esta convención la rama del lado derecho no tiene vér tices donde hemos señalado puntos en los niveles 0 a 3 , y en consecuencia constituye una sola rama y no varias , por lo cual los mencionados puntos pueden suprimirse .

Cada conjunto de N elementos tendrá un árbol binario característico si se adoptan las convenciones de selección que hemos enunciado . En la tabla 1 se ilustran algunos ejemplos .

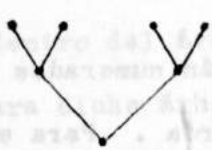
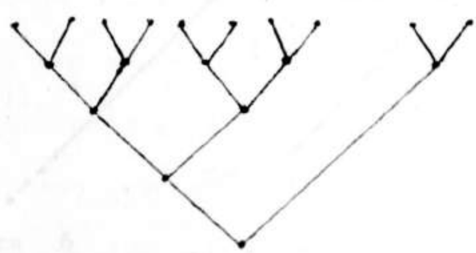
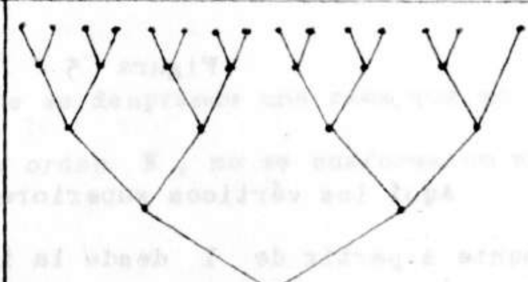
N	ARBOL BINARIO
1	
2	
3	
4	
10	
15	

TABLA No. 1

4. ARBOL BINARIO INFINITO

Podemos considerar un árbol de orden N como una porción de un árbol binario infinito, por ejemplo, el de la figura 5.

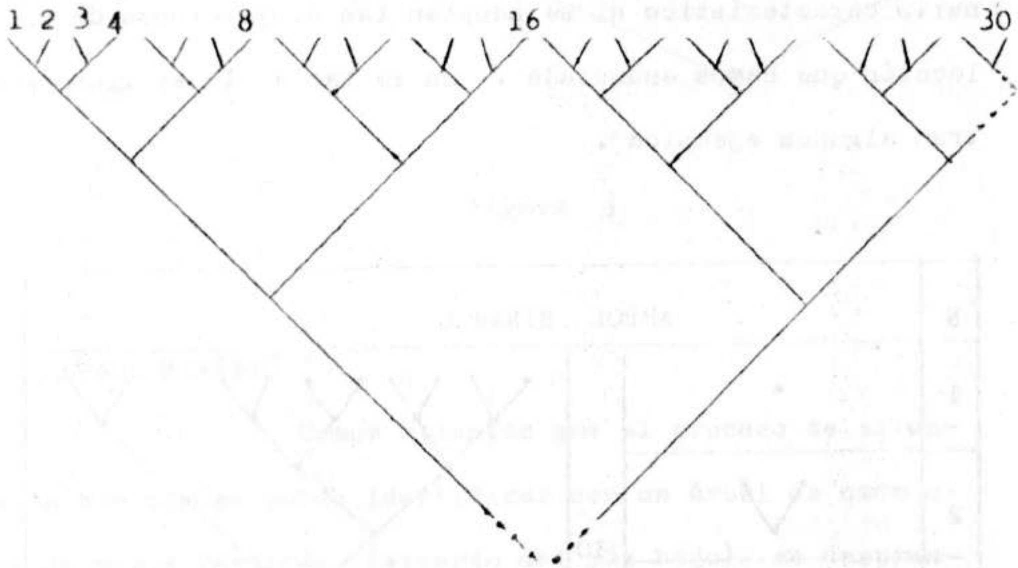


Figura 5

Aquí los vértices superiores están numerados sucesivamente a partir de 1 desde la izquierda. Para encontrar el árbol que le pertenece a un número dado, se comienza por el vértice superior que le corresponde a dicho número y se va descendiendo hasta encontrar la rama extrema de la izquierda . Todas las ramas y todos los vértices que están comprendidos entre esta rama izquierda y la recorrida por el procedimiento anterior conforman el árbol de orden N .

Ejemplo : N = 6

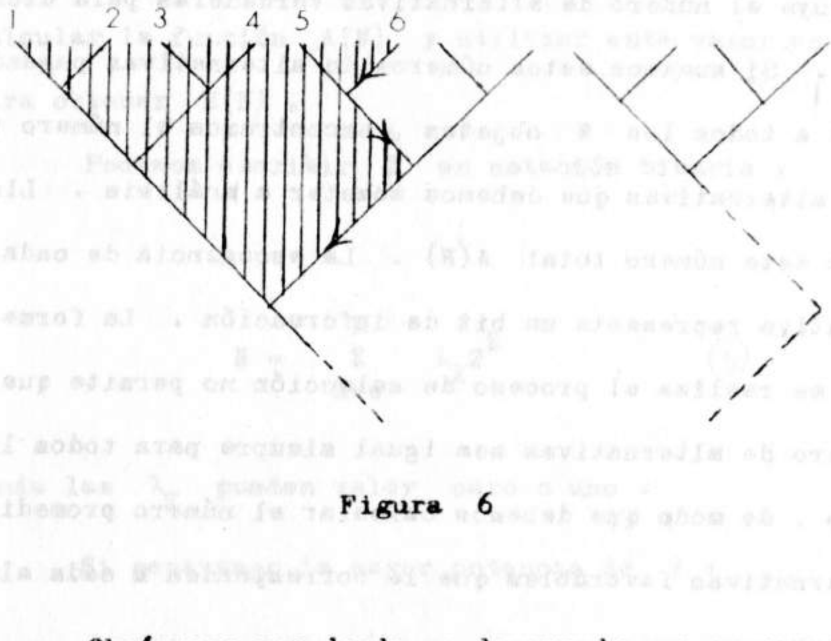


Figura 6

Obsérvese que donde se desprende una rama que no queda dentro del árbol de orden N , no se conforma un vértice para dicho árbol .

5. MEDICION DE LA INFORMACION POR ALTERNATIVAS

La información total en un proceso de selección de uno entre N eventos (o elementos , u objetos) igualmente probables se puede calcular en bits contando el número de alternativas binarias que se presentan en el proceso de selección . Se obtiene información en la medida en que se eliminan todas las alternativas falsas y se opta por las verdaderas . El número de ramas que encontramos desde el vértice inferior del árbol hasta el objeto seleccionado , constituye el número de alternativas verdaderas para dicho objeto . Si sumamos estos números de alternativas pasando revista a todos los N objetos , encontramos el número total de alternativas que debemos someter a análisis . Llamemos a este número total $A(N)$. La escogencia de cada alternativa representa un bit de información . La forma en que se realiza el proceso de selección no permite que el número de alternativas sea igual siempre para todos los objetos , de modo que debemos calcular el número promedio de alternativas favorables que le corresponden a cada elemento como

$$H(N) = \frac{A(N)}{N} \qquad (4)$$

Esta será la información promedio asociada con la selección de uno entre N elementos .

Vamos a mostrar que $H(N)$ goza de propiedades muy semejantes a las de la función $I(N)$ definida en (3), y coincide con ella cuando es posible escribir $N = 2^M$, con M un número natural.

6. CALCULO DE $H(N)$

Realmente, nuestro problema es calcular la función $A(N)$ y utilizar este valor en (4) para obtener $H(N)$.

Podemos escribir N en notación binaria :

$$N = \sum_{K=0}^{M} \lambda_K 2^K \tag{5}$$

donde las λ_K pueden valer cero o uno.

Si separamos la mayor potencia de 2 :

$$N = \sum_{K=0}^{M-1} \lambda_K 2^K + 2^M = 2^M + Q \tag{6}$$

pues $\lambda_M = 1$ necesariamente, obtenemos

$$A(N) = A(2^M + Q) . \tag{7}$$

donde $Q < 2^M$.

Considerando el árbol binario, el árbol de 2^M+Q se compone de dos ramas principales que parten de un vértice : la rama de la izquierda conduce al vértice inferior del árbol de 2^M, mientras que la rama derecha conduce al vértice inferior del árbol correspondiente a Q. El árbol del conjunto contiene entonces dos ramas más que los sub-árboles por separado. Estas ramas adicionales son las que se unen en el vértice inferior. De esta manera, cada elemento del conjunto total tiene que ser localizado después de haber agregado una alternativa más. Por lo tanto, hemos agregado en total, N alternativas en comparación con la situación en la cual 2^M y Q están separados. Encontramos como consecuencia de este razonamiento :

$$A(2^M + Q) = A(2^M) + A(Q) + N \qquad (8)$$

Donde $A(2^M)$ y $A(Q)$ son los números de alternativas correspondientes a 2^M y Q respectivamente, tomado cada conjunto por separado. Asumimos $A(0) = 0$.

La fórmula (8) no es válida si $Q = 0$, como se puede ver reemplazando directamente. Esto quiere decir que debemos calcular $A(2^M)$ por separado para poder obtener de (8) una fórmula de recurrencia que nos con duzca al resultado final .

En el árbol de cualquier potencia natural de 2 , el número de ramas para llegar desde el vértice inferior a uno de los elementos es el mismo para todos ellos , y es precisamente igual a dicha potencia de 2 ; de modo que el número total de alternativas para 2^M elementos es

$$A(2^M) = M\, 2^M \qquad (9)$$

La fórmula (9) nos permite calcular H directamente cuando $N = 2^M$:

$$H(2^M) = \frac{A(2^M)}{2^M} = M \qquad (10)$$

con lo cual , según (1) , resulta que $H(2^M) = I(2^M)$. Es decir , la función H coincide con la función logarítmica de información cuando N es una potencia natural de 2 .

Ahora , según (6)

$$Q = \sum_{K=0}^{M-1} \lambda_K\, 2^K \qquad (11)$$

donde λ_{M-1} puede tomar valor cero o uno .

Si $\lambda_{M-1} = 0$, entonces

$$A(Q) = A\left(\sum_{K=0}^{M-2} \lambda_K 2^K\right) \qquad (12)$$

Si $\lambda_{M-1} = 1$, entonces

$$A(Q) = A\left(2^{M-1} + \sum_{K=0}^{M-2} \lambda_K 2^K\right) ,$$

y si utilizamos la fórmula (8) reemplazando adecuadamente :

$$A(Q) = A(2^{M-1}) + A\left(\sum_{K=0}^{M-2} \lambda_K 2^K\right) + Q \qquad (13)$$

Las fórmulas (12) y (13) se pueden reducir a una sola si utilizamos λ_{M-1} :

$$A(Q) = \lambda_{M-1} \left[A(2^{M-1}) + Q\right] + A\left(\sum_{K=0}^{M-2} \lambda_K 2^K\right) \qquad (14a)$$

pues si $\lambda_{M-1} = 0$, (14a) se reduce a (12) y si $\lambda_{M-1} = 1$, se convierte en (13) .

Si sustituímos Q por su valor , encontramos

$$A\left(\sum_{K=0}^{M-1}\lambda_K 2^K\right)=A\left(\sum_{K=0}^{M-2}\lambda_K 2^K\right)+\lambda_{M-1}\left[A(2^{M-1})+\sum_{K=0}^{M-1}\lambda_K 2^K\right] \quad (14b)$$

que es la fórmula de recurrencia buscada .

Utilizando (14b) sucesivamente , obtenemos la secuencia de ecuaciones

$$A\left(\sum_{K=0}^{M}\lambda_K 2^K\right)=A\left(\sum_{K=0}^{M-1}\lambda_K 2^K\right)+\lambda_M\left[A(2^M)+\sum_{K=0}^{M}\lambda_K 2^K\right]$$

$$A\left(\sum_{K=0}^{M-1}\lambda_K 2^K\right)=A\left(\sum_{K=0}^{M-2}\lambda_K 2^K\right)+\lambda_{M-1}\left[A(2^{M-1})+\sum_{K=0}^{M-1}\lambda_K 2^K\right]$$

$$A\left(\sum_{K=0}^{M-2}\lambda_K 2^K\right)=\ldots\ldots$$

$$\vdots$$

$$A\left(\sum_{K=0}^{1}\lambda_K 2^K\right)=A(\lambda_o)+\lambda_1\left[A(2)+\sum_{K=0}^{1}\lambda_K 2^K\right]$$

Escribimos lo anterior con la precaución de que la última ecuación aparece sólo en caso de que ninguna de las sumas anteriores $\sum_{K=0}^{P}\lambda_K 2^K$ sea exactamente una potencia de 2 . Es decir , si

$$\sum_{K=0}^{P}\lambda_K 2^K=2^P \quad (15)$$

o sea , si $\lambda_P = 1$ y $\lambda_K = 0$ con $K < P$, entonces la última ecuación de la secuencia es

$$A\left(\sum_{K=0}^{P} \lambda_K 2^K\right) = A(2^P) = P2^P$$

de acuerdo con (9) .

Si sumamos las ecuaciones de la secuencia miembro a miembro , todos los términos de la izquierda , excepto el primero , se cancelan y obtenemos (recordando que $A(\lambda_o) = 0$ dado que λ_o vale cero o uno) :

$$A\left(\sum_{K=0}^{M} \lambda_K 2^K\right) = \sum_{j=1}^{M} \lambda_j \left[A(2^j) + \sum_{K=0}^{j} \lambda_K 2^K \right] . \quad (16)$$

En caso de que $\lambda_K = 0$ para $K < P$, el resultado de la suma es

$$A\left(\sum_{K=0}^{M} \lambda_K 2^K\right) = \sum_{j=P+1}^{M} \lambda_j \left[A(2^j) + \sum_{K=0}^{j} \lambda_K 2^K \right] + A(2^P) \quad (17)$$

y esta fórmula evidentemente se reduce a (16) cuando $P = 0$.

Reorganizando un poco , podemos escribir

$$A\left(\sum_{K=0}^{M} \lambda_K 2^K\right) = \sum_{j=P}^{M} j\lambda_j 2^j + \sum_{j=P+1}^{M} \lambda_j \sum_{K=0}^{j} \lambda_K 2^K \quad (18)$$

donde P cumple la condición (15) .

Esta es , finalmente , la fórmula general para $A(N)$ cuando escribimos N en la forma (5) . Si $N=2^M$, o sea , si $P = M$, el segundo término de (18) pierde sentido y el primero vale $M\,2^M$, lo cual está de acuerdo con (9) .

7. RESULTADOS NUMERICOS

a) Ejemplo : calculemos $A(N)$ donde $N = 13$. En el sistema binario

$$N = 1101 .$$

Aquí $P = 0$.

Podríamos hacer $\lambda_o = 1$, $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 1$ y reemplazar en (18) , pero existe una interpretación más cómoda para cálculos manuales : los términos de la primera sumatoria de (18) pertenecen a la sucesión

$$\left\{ j2^j \right\} = \left\{ 0,2,18,24,64,160,384,896,\ldots \right\} \qquad (19)$$

y cada uno contribuye o no según que la correspondiente ci fra binaria de N valga uno o cero (contadas de derecha a izquierda y a partir de cero) .

En este ejemplo contribuyen 8 y 24 cuya suma es 32 .

Los términos de la segunda sumatoria de (18) representan los números binarios que van apareciendo cuando se toman sucesivamente las dos últimas cifras de N , las tres últimas cifras de N , etc. Solo se toman en cuenta los números binarios que comienzan por 1 , exceptuando el que esté seguido únicamente por ceros . En este caso van apareciendo :

2 últimas cifras	01	no se toma en cuenta
3 últimas cifras	101 = 5	
4 últimas cifras	1101 =13	
	Suma =18 .	

Entonces , A(13) = 32 + 18 = 50 .

b) Ejemplo : calcular A(N) con N = 88 . En binario :

N = 1011000 .

Aquí P = 3 .

Recurriendo al método desarrollado en el ejemplo anterior , al primer término de (18) contribuyen los siguientes números de la sucesión (19) :

$$24 + 64 + 384 = 472 .$$

Al segundo término contribuyen

$$11000 = 24$$

$$y \quad 1011000 = 88$$

$$\text{suma} \quad = 112 .$$

Entonces $A(88) = 472 + 112 = 584$.

A continuación se tabulan para $1 \leq N \leq 100$ los valores de $A(N)$, $H(N)$, $\log_2 N$ y de la función

$$D(N) = H(N) - \log_2 N .$$

Estos valores fueron calculados con un minicomputador HP9830B .

Las gráficas de $H(N)$ y $D(N)$ se presentan en las figuras 7 y 8 .

<u>TABLA</u> 2

N	A(N)	H(N)	LOG_2N	DIF=D(N)
2	2	1.000000	1.000000	0.000000
3	5	1.666667	1.584963	0.081704
4	8	2.000000	2.000000	0.000000
5	13	2.600000	2.321923	0.278072
6	16	2.666667	2.584963	0.081704

N	A(N)	H(N)	LOG_2N	DIF=D(N)
7	20	2.857143	2.807355	0.049788
8	24	3.000000	3.000000	0.000000
9	33	3.666667	3.169925	0.496742
10	36	3.600000	3.321923	0.278072
11	40	3.636364	3.459432	0.176932
12	44	3.666667	3.584963	0.081704
13	50	3.846154	3.700440	0.145714
14	54	3.857143	3.807355	0.049788
15	59	3.933333	3.906891	0.026443
16	64	4.000000	4.000000	0.000000
17	31	4.764706	4.087463	0.677243
18	84	4.666667	4.169925	0.496742
19	88	4.631579	4.247928	0.383651
20	92	4.600000	4.321928	0.278072
21	98	4.666667	4.392317	0.274349
22	102	4.636364	4.459432	0.176932
23	107	4.652174	4.523562	0.128612
24	112	4.666667	4.584963	0.081704
25	122	4.880000	4.648856	0.236144
26	126	4.846154	4.700440	0.145714
27	131	4.851852	4.754888	0.096964

N	A(N)	H(N)	LOG_2N	DIF-D(N)
28	136	4.857143	4.807355	0.049788
29	143	4.931034	4.857981	0.073053
30	148	4.933333	4.906891	0.026442
31	154	4.967742	4.954196	0.013546
32	160	5.000000	5.000000	0.000000
33	193	5.848485	5.044394	0.804091
34	196	5.764706	5.087463	0.677243
35	200	5.714286	5.129283	0.585003
36	204	5.666667	5.169925	0.496742
37	210	5.675676	5.209453	0.466222
38	214	5.631579	5.247928	0.383651
39	219	5.615385	5.285402	0.329982
40	224	5.600000	5.321928	0.278072
41	234	5.707317	5.357552	0.349765
42	238	5.666667	5.392317	0.274349
43	243	5.651163	5.426265	0.224898
44	248	5.636364	5.459432	0.176932
45	255	5.666667	5.491853	0.174814
46	260	5.652174	5.523562	0.128612
47	266	5.659574	5.554589	0.104986
48	272	5.666667	5.584963	0.081704

N	A(N)	H(N)	$LOG_2 N$	DIF-D(N)
49	290	5.918367	5.614710	0.303658
50	294	5.880000	5.643856	0.236144
51	299	5.862745	5.672425	0.190320
52	304	5.846154	5.700440	0.145714
53	311	5.867925	5.727920	0.140004
54	316	5.851852	5.754888	0.096964
55	322	5.854545	5.781360	0.073186
56	328	5.857143	5.807355	0.049788
57	339	5.947368	5.832890	0.114478
58	344	5.931034	5.857931	0.073053
59	350	5.932203	5.882643	0.049560
60	356	5.933333	5.906891	0.026443
61	364	5.967213	5.930737	0.036476
62	370	5.967742	5.954196	0.013546
63	377	5.984127	5.977280	0.006847
64	384	6.000000	6.000000	0.000000
65	449	6.907692	6.022368	0.885324
66	452	6.848485	6.044394	0.804091
67	456	6.805970	6.066089	0.739881
68	460	6.764706	6.087463	0.677243
69	466	6.753623	6.108524	0.645099

N	A(N)	H(N)	LOG_2N	DIF-D(N)
70	470	6.714286	6.129283	0.585003
71	475	6.690141	6.149747	0.540394
72	480	6.666667	6.169925	0.496742
73	490	6.712329	6.189825	0.522504
74	494	6.675676	6.209453	0.466222
75	499	6.653333	6.228819	0.424515
76	504	6.681579	6.247928	0.333651
77	511	6.636364	6.266787	0.369577
78	516	6.615385	6.285402	0.329982
79	522	6.607595	6.303781	0.303814
80	528	6.600000	6.321928	0.278072
81	546	6.740741	6.330850	0.400391
82	550	6.707317	6.357552	0.349765
83	555	6.686747	6.375039	0.311708
84	560	6.666667	6.392317	0.274349
85	567	6.670588	6.409391	0.261197
86	572	6.651163	6.426265	0.224898
87	578	6.643678	6.442943	0.200735
88	584	6.636364	6.459432	0.176932
89	595	6.685393	6.475733	0.209660
90	600	6.666667	6.491853	0.174814

N	A(N)	H(N)	LOG_2N	DIF=D(N)
91	606	6.659341	6.507795	0.151546
92	612	6.652174	6.523562	0.128612
93	620	6.666667	6.539159	0.127508
94	626	6.659574	6.554589	0.104986
95	633	6.663158	6.569856	0.093302
96	640	6.666667	6.584963	0.081704
97	674	6.948454	6.599913	0.348541
98	678	6.918367	6.614710	0.303658
99	683	6.898990	6.629357	0.269633
100	688	6.880000	6.643856	0.236144

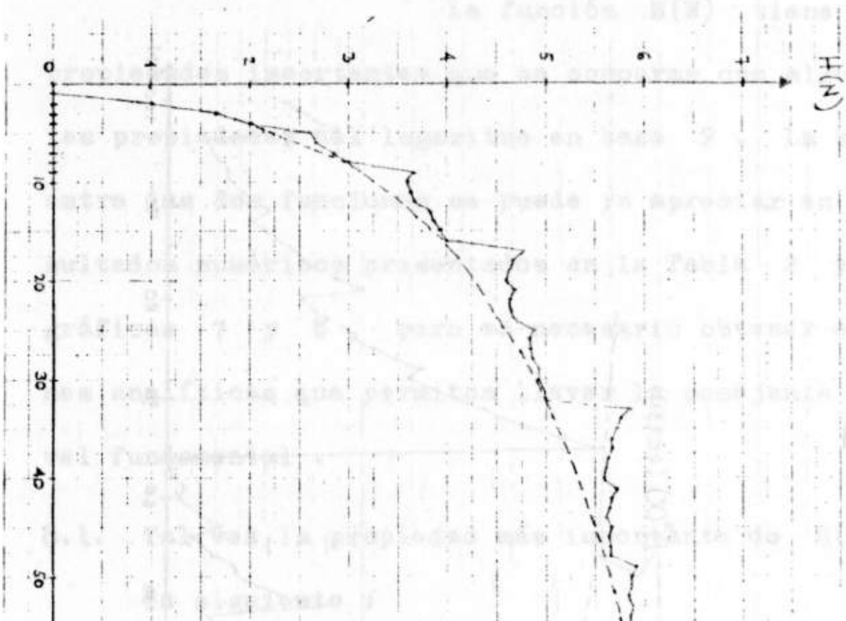


Figura 7

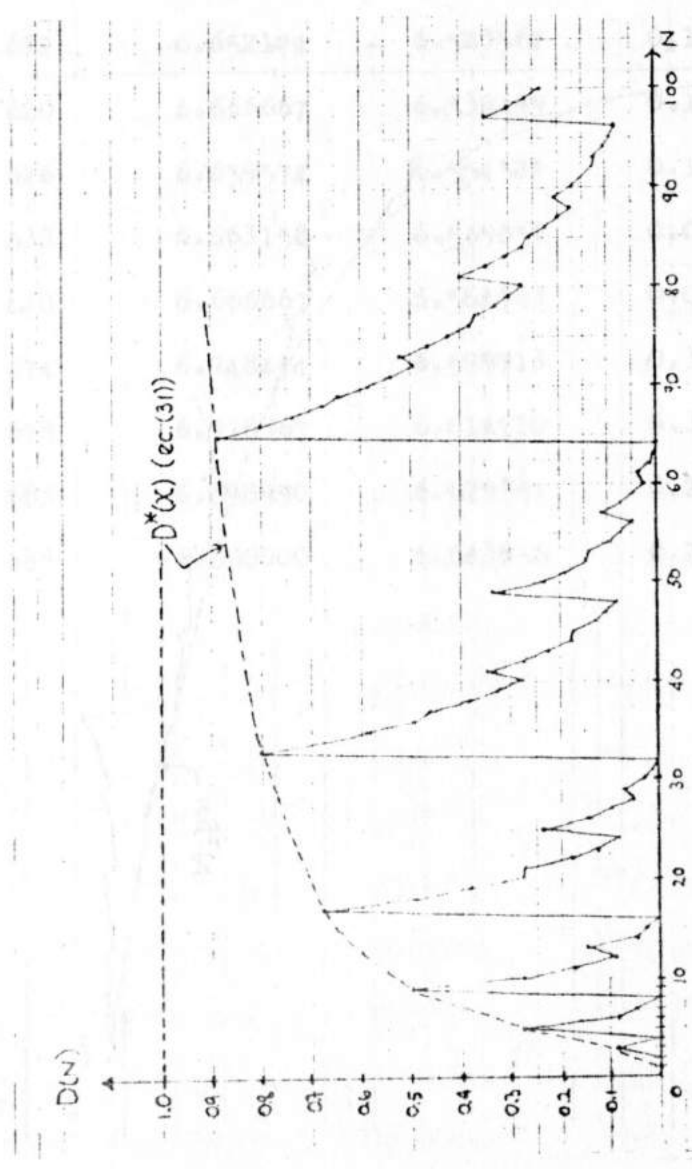

D(N)
D*(X) (ec.(31))
N
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50
60
70
80
90
100

Figura 8

8. PROPIEDADES DE $H(N)$

La función $H(N)$ tiene varias propiedades importantes que se comparan con algunas de las propiedades del logaritmo en base 2 . La semejanza entre las dos funciones se puede ya apreciar en los resultados numéricos presentados en la Tabla 2 y en las gráficas 7 y 8 , pero es necesario obtener expresiones analíticas que permitan llevar la semejanza a un nivel fundamental .

8.1. Tal vez la propiedad más importante de $H(N)$ es la siguiente :

$$H(2N) = H(N) + 1 \tag{20}$$

que, como se ve , es enteramente semejante a la del logaritmo :

$$\log_2 2N = \log_2 N + 1 . \tag{21}$$

De aquí se desprende el que llamemos a $H(N)$ una función cuasi-logarítmica .

La demostración es la siguiente . Es claro que si escribimos N en forma binaria

$$N = \sum_{K=0}^{M} \lambda_K 2^K ,$$

entonces ,

$$2N = 2 \sum_{K=0}^{M} \lambda_K 2^K = \sum_{K=0}^{M} \lambda_K 2^{K+1} ,$$

o , si cambiamos $K + 1$ por n ,

$$2N = \sum_{n=1}^{M+1} \lambda_{n-1} 2^n = \sum_{n=0}^{M+1} \delta_n 2^n , \quad \text{con} \quad \delta_o = 0 , \tag{22}$$

o sea , que el número binario que se obtiene es el mismo anterior con un cero agregado a la derecha , o , lo que es lo mismo ,

$$\delta_o = 0 , \quad \delta_n = \lambda_{n-1} \quad (n = 1,\dots, M + 1) . \tag{23}$$

Si utilizamos (18) para la expresión (22) , obtenemos

$$A\left(\sum_{n=0}^{M+1} \delta_n 2^n \right) = \sum_{n=q}^{M+1} \delta_n n 2^n + \sum_{j=q+1}^{M+1} \delta_j \sum_{n=0}^{j} \delta_n 2^n , \tag{24}$$

donde es claro que $q = p+1$ pues $2N$ tiene los mismos dígitos que N , sólo que desplazados a la izquierda un lugar . Si hacemos en (24) nuevamente $n = K+1$,

$$A(2N) = \sum_{K=q-1}^{M} \delta_{K+1}(K+1)2^{K+1} + \sum_{j=q+1}^{M+1} \delta_j \sum_{K=0}^{j-1} \delta_{K+1} 2^{K+1}$$

y si transformamos $j = e+1$ y hacemos aparecer p en lugar de q ,

$$A(2N) = \sum_{K=p}^{M} \delta_{K+1}(K+1)2^{K+1} + \sum_{e=p+1}^{M} \delta_{e+1} \sum_{K=0}^{e} \delta_{K+1} 2^{K+1} ,$$

de modo que si utilizamos (23) , obtenemos

$$A(2N) = \sum_{K=p}^{M} \lambda_K (K+1)2^{K+1} + \sum_{e=p+1}^{M} \lambda_e \sum_{K=0}^{e} \lambda_K 2^{K+1} ,$$

en donde podemos factorizar un 2 y separar la primera sumatoria , lo cual nos da

$$A(2N) = 2 \left[\sum_{K=p}^{M} \lambda_K 2^K + \sum_{K=p}^{M} \lambda_K K 2^K + \sum_{e=p+1}^{M} \lambda_e \sum_{K=0}^{e} \lambda_K 2^K \right]$$

y , finalmente ,

$$A(2N) = 2 \left[N + A(N) \right] , \tag{25}$$

pues

$$\sum_{K=p}^{M} \lambda_K 2^K = \sum_{K=0}^{M} \lambda_K 2^K = N ,$$

y los dos últimos sumandos dentro del paréntesis coinciden con la expresión (18) .

Si dividimos (25) por 2N a ambos lados , encontramos

$$H(2N) = \frac{A(2N)}{2N} = 1 + \frac{A(N)}{N}$$

Es decir ,

$$H(2N) = H(N) + 1 \quad , \tag{20}$$

tal como queríamos demostrar .

8.2. Como corolario de la propiedad expresada en (20) , es inmediato que se puede demostrar por inducción la siguiente :

$$H(2^M N) = H(N) + M \quad ; \tag{26}$$

de modo que para $N = 1$ y recordando que $H(1) = 0$, se obtiene

$$H(2^M) = M \quad ,$$

que es la expresión (10) hallada al principio .

8.3. Otra propiedad que es facilmente obtenible y notable por su forma se obtiene a partir de (8) :

$$A(N) = A(2^M+Q) = A(2^M)+A(Q)+N \quad ,$$

y si dividimos por $N = 2^M + Q$,

$$H(2^M+Q) = \frac{A(2^M)+A(Q)}{2^M+Q} + 1$$

Pero $A(2^M) = 2^M H(2^M)$ y $A(Q) = QH(Q)$, de donde

$$H(2^M+Q) = \frac{2^M H(2^M)+QH(Q)}{2^M+1} + 1 , \qquad (27)$$

cuyo primer término tiene la forma de un promedio ponderado . Es importante recordar que en esta fórmula se debe tener $Q < 2^M$, de acuerdo con el razonamiento que conduce a la ecuación (6) .

8.4. Observando la gráfica de $D(N)$ se encuentra a simple vista que la función tiene una serie de máximos relativos en los puntos expresables en la forma $N = 2^M + 1$. Más exactamente , entre $N_0 = 2^M$ y $N_1 = 2^{M+1}$, la función es máxima en $N = 2^M + 1$; o sea , que si $1 < R < 2^M$, entonces

$$D(2^M + 1) > D(2^M + R) . \qquad (28)$$

La demostración de (28) procede así :

$D(2^M+1) = H(2^M+1) - Log_2(2^M+1)$, por definición .

$$= \frac{2^M\ H(2^M)+H(1)}{2^M+1} + 1 - \log_2(2^M+1)$$, por aplicación de (27) .

Entonces

$$D(2^M+1) = 1 + \frac{M2^M}{2^M+1} - \left[\log_2\left(\frac{2^M+1}{2^M}\right) + \log_2 2^M\right]$$

$$= 1 + \frac{M2^M}{2^M+1} - \log_2\left(1+\frac{1}{2^M}\right) - M \ .$$

Finalmente ,

$$D(2^M+1) = 1 - \frac{M}{2^M+1} - \log_2\left(1+\frac{1}{2^M}\right) \ . \qquad (29)$$

Este resultado es de suma importancia , como veremos más adelante en el parágrafo 8.6.

De la misma manera se puede obtener

$$D(2^M+R) = 1 - \frac{R\ M}{2^M+R} - \log_2\left(1+\frac{R}{2^M}\right) \qquad (30)$$

con $R < 2^M$, según la aclaración que se hace a continuación de la ecuación (27) .

De aquí es fácil ver por comparación de (29) y (30) que la ecuación (28) se cumple si $R > 1$.

8.5. En la gráfica de $D(N)$ se puede ver también que los máximos que aparecen en $N = 2^M + 1$ son cada vez más pronunciados , de modo que es de esperarse que $D(2^M + 1)$ sea una función creciente con respecto a la variable M .

Sea $x = 2^M + 1$ ó $M = \log_2(x - 1)$, o sea que se puede escribir

$$D(x) = 1 - \frac{\log_2(x-1)}{x} - \log_2\left(\frac{x}{x-1}\right)$$

$$= 1 + \frac{(x-1)}{x} \log_2(x-1) - \log_2 x \, .$$

Esta función sólo está definida en los puntos $x = 2^M + 1$, pero si tomamos a X como una variable contínua , tendremos una función definida para todo $X > 1$:

$$D^{\sim}(X) = 1 + \frac{X-1}{X} \log_2 (X-1) - \log_2 X \, , \qquad (31)$$

cuya derivada es

$$\frac{d}{dX} D^{\sim}(X) = \frac{\log_2(X-1)}{X^2} \geq 0 \qquad (X \geq 2) \, ,$$

o sea que $D^{\sim}(X)$ es creciente , y por lo tanto también $D(2^M+1)$ lo es con respecto a la variable M .

8.6. Teniendo en cuenta el último resultado sobre la monotonía de $D(2^M + 1)$ y la ecuación (29), podemos fácilmente encontrar

$$\lim_{M \to \infty} D(2^M + 1) = 1 . \tag{32}$$

Esta expresión indica que la máxima diferencia entre $H(N)$ y $\log_2 N$ tiende a 1 . De este modo , aunque la diferencia oscila y tiene máximos cada vez mayores , la función $H(N)$ permanece siempre en la vecindad de $\log_2 N$. Esta es otra razón para llamar a $H(N)$ una función cuasi-logarítmica .

8.7. Utilizando las ecuaciones (20) y (21) , se obtiene en forma inmediata

$$D(2N) = D(N) , \tag{33}$$

expresión que pone de manifiesto una peculiar forma de " periodicidad " de la función $D(N)$ a pesar de su aparente oscilar caprichoso . Esta ecuación puede visualizarse fácilmente en la tabla 2 ó en la figura 8 , escogiendo cualquier número (en este caso menor o igual

que 50) , y multiplicándolo por 2 .

8.8. No hemos logrado aún demostrar una propiedad que parece evidente de los cálculos realizados y de la figura 8 , pero la proponemos como una conjetura :

$$D(N) \geq 0 \quad . \tag{34}$$

9. CONCLUSIONES

Al utilizar el método binario de selección descrito , esperamos obtener una información expresada por la función $H(N)$. Sinembargo , puesto que $H(N)$ no es creciente con respecto a N , no satisface una de las mínimas condiciones exigibles a la información que debe asociarse con la escogencia de uno entre N eventos .

Si la conjetura (34) es válida , esto significa que la información esperada es mayor que la información que podríamos llamar intrínseca (es decir , la proporcionada por la fórmula de Shannon) . De aquí deduciríamos que el proceso de selección escogido no es eficaz en un cien por ciento , excepto cuando $N = 2^M$, en cuyo caso $D(N) = 0$, como ya se demostró .

Por supuesto , hace falta definir una medida de " eficiencia " , motivo de una futura investigación .

Bástenos señalar , por ejemplo , que en este caso la "ineficacia " se hace protuberante cuando $N = 2^M + 1$. Este resultado es de esperarse intuitivamente , pues al agregarle un evento a un conjunto que ya contiene 2^M , estamos empleando en forma mínima las posibilidades de obtener la información que brinda la aparición de un nuevo nivel de alternativas . Para ilustrar esto , pensemos en la poca economía que representaría agregar a una calculadora electrónica un lugar decimal adicional que sólo va a ser ocupado por los dígitos cero ó uno .

De las últimas observaciones puede inferirse que el método desarrollado aquí y sus resultados quizá tengan utilidad en el análisis de decisiones , en la formulación de códigos , y en otras aplicaciones donde se requiera utilizar algún criterio de economía o de enumeración exacta de alternativas . [3] , [4] .

Finalmente , vale la pena pensar en distintas extensiones que los conceptos aquí formulados podrían recibir. La más evidente sería el cálculo de funciones H(N) que correspondieran a procesos de selección ternarios , cuaternarios , o en general , de orden n , con posible abstracción a un orden no necesariamente entero .

Por otra parte , tomando como punto de partida la ecuación (20) que establece la propiedad cuasi-logarítmica , es interesante pensar en la generalización del concepto de función cuasi-logarítmica , lo cual podría talvez (y ojalá) coincidir con la extensión a los otros procesos de selección señalados arriba .

Por último , es natural pensar en una generalización de otro orden , que parta de la asignación de probabilidad no uniforme a los diferentes eventos objeto de la selección .

10. AGRADECIMIENTOS

Deseo agradecer al profesor Jorge Rodríguez del Departamento de Matemáticas de la Universidad del Valle , sus múltiples sugerencias con respecto a una buena parte del manuscrito de este trabajo .

BIBLIOGRAFIA

1. Shannon, C.E. and Weaver, W. The Mathematical Theory of Communication. University of Illinois Press, Urbana, 1949 .

2. Khinchin, A.I. Mathematical Foundations of Informa-

tion Theory. Dover Publications, Inc., New York, 1957.

3. Pierce, J.R. **Symbols, Signals and Noise.** Harper y Row, Publishers, New York , 1965 .

4. Bell, D.A. **Information Theory and its Engineering Aplications.** Sir Isaac Pitman y Sons, Ltd., London 1968 .

Jaime HERNANDEZ G.

Departamento de Matemáticas

Universidad Tecnológica de Pereira

Pereira - Colombia .