UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Centro Universitario de Investigaciones Bibliotecológicas

Seminario Hispanoamericano sobre Estudios Métricos de la Información

BRADFORD Y MARKOV: dos casos de aplicación a bibliotecas mexicanas

Alvaro Quijano Solís,
Heshmatallah Khorramzadeh,
Biblioteca Daniel Cosío Villegas
El Colegio de México, A.C.

Agosto 28-Septiembre 1o., 1995

INTRODUCCION

Cuando pensamos en el método científico, nuestra mente asocia la experimentación como una parte esencial de él. En sentido estricto, la experimentación se refiere a la manipulación física de variables. Este concepto está lejos de la realidad de la administración de bibliotecas. En ellas, la toma de decisiones no siempre puede tomar en consideración el número infinito de variables que forman parte de su entorno y, mucho menos, experimentar con diferentes soluciones que afecten a toda la organización. Podemos, sin embargo, observar el sistema, aunque no lo manipulemos. Podemos construir representaciones del sistema en observación y de sus operaciones y, sobre las primeras, conducir investigación y experimentación para facilitar el acercamiento a los objetivos de la organización. Podemos, en suma, construir modelos

Una situación real en la práctica bibliotecológica puede involucrar un número sustancial de variables. Afortunadamente, la experiencia nos muestra que una fracción de estas variables domina la mayor parte del comportamiento del sistema. Por ende, la simplificación del sistema real en términos de un modelo se concentra principalmente en la identificación de las variables más importantes y las principales relaciones que lo gobiernan.

Puede decirse que la esencia de un enfoque científico al proceso de toma de decisiones es la construcción de modelos. Los modelos más útiles son aquellos que representan las propiedades más relevantes de una realidad y, por lo general, son matemáticos en su naturaleza. En general, se definen y se miden las variables importantes de un sistema, y la observación de la realidad y la intuición gerencial se combinan para establecer relaciones dominantes. Estas se utilizan con el fin de describir y explicar el pasado, ejercitar el control en el presente y predecir el futuro.

Los modelos pueden usarse para encontrar el impacto que las diferentes alternativas tienen sobre el desempeño de la organización y para escoger las más adecuadas o que producen el mayor beneficio. Con el fin de lograr estos resultados, un modelo debe contener cinco elementos básicos:

1. Medida de desempeño que averigüe el grado con el que los objetivos se logran (función objetivo)

2. Variables que son sujetas a control del administrador (variables de decisión o alternativas)

3. Factores que afectan el desempeño, pero que no son sujetos a control del administrador (variables no controlables o parámetros)

4. Relación funcional entre las variables controlables, los parámetros y la medida de desempeño.

5. Especificación de restricciones sobre las variables controlables

Una vez construido el modelo, puede utilizarse para encontrar, exacta o aproximadamente, los valores óptimos de las variables controlables y que son los que deberán producir el mejor desempeño del sistema para un valor especifico de los parámetros; es decir, podemos encontrar una solución al problema, a partir del modelo. Las soluciones pueden encontrarse por simulación o experimentación sobre el modelo o pueden encontrarse por métodos analíticos, generalmente matemáticos.

En general, puede decirse que modelar la realidad confronta dos condiciones que, en ocasiones, suelen ser conflictivas entre ellas: hacer un modelo fácil de resolver y que, al mismo tiempo, sea tan preciso como la realidad misma. La regla práctica podría ser que un modelo aproximado de un sistema, que mejora el desempeño de éste, es mejor que un modelo exacto que no lo logra. Demasiada complejidad inhibe la derivación de soluciones relevantes para la toma de decisiones.

Una característica útil de un modelo de decisión es que indica qué datos son relevantes para el análisis de las decisiones alternativas. A este respecto, los modelos son útiles para diseñar los sistemas de datos. Sin embargo, no es cierto concluir que los modelos determinan, por sí mismos, qué datos deben ser recabados. Es decir, debe buscarse el balance entre la recolección de datos ideales y el uso de datos disponibles. Los modelos deben modificarse para aprovechar los datos disponibles.

El enfoque científico de la toma de decisiones no elimina el uso de la intuición administrativa. Una combinación efectiva de modelos de decisiones e intuición dará como resultado mejores decisiones y un nivel más alto de liderazgo. Con este, el administrador podrá plantear preguntas del tipo "qué pasaría si..." en las cuales hipotetizará la realización de varias alternativas de decisión y evaluará los efectos sobre el desempeño de la organización.

Los métodos cuantitativos, por sí mismos, representan siempre una visión importante, pero limitada, del comportamiento de los sistemas bibliotecarios y de sus usuarios. La experiencia muestra que aquéllos deben enriquecerse con el análisis situacional (por ejemplo, la metodología de sistemas blandos) y con la experiencia profesional.

En esta ponencia se resumen dos de los modelos cuantitativos más importantes en el campo de bibliotecología y ciencias de la información, que han sido sujetos a extensos debates y a controversia científica en una gama de artículos, libros y comentarios. Estos dos modelos son: la ley de Bradford y sus modificaciones y el Modelo Morse-Markov y su extensión. Se presentan estos dos modelos según su aplicación en situaciones concretas. Se hace un esfuerzo para explicarlos en una forma comprensible, con la finalidad de que algunos problemas que se presentan en las bibliotecas puedan ser considerados con enfoques científicos que contribuyan a su solución.

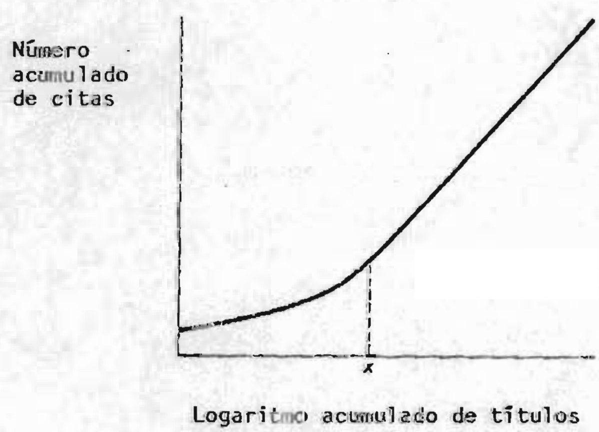
LEY DE BRADFORD

En los años 30 mientras Bradford buscaba artículos sobre Geofísica aplicada y lubricación, llegó a descubrir un patrón común en la disperción de estos artículos en revistas científicas. Con el análisis de estos dos grupos de artículos publicó los resultados, primero en un artículo en 1934 y, posteriormente, en un libro titulado Documentation en 1948.

Bradford arregló las 326 revistas que contenían 1332 artículos sobre geofísica aplicada en orden decreciente en relación al número de artículos de cada revista y llevó a cabo el mismo procedimiento para las 164 revistas que tenían 395 artículos sobre lubricación. En ambos casos, observó tres patrones: 1) unas cuantas revistas con un gran número de artículos, un grupo más numeroso de revistas con pocos artículos y muchas revistas con solamente un artículo.

Bradford ordenó las revistas de acuerdo con el número de artículos en cada una; en otras palabras la revista con mayor número de artículos tenía la posición uno y así sucesivamente. Después, realizó dos operaciones mas; tomó el logaritmo de la suma acumulativa de revistas en orden descendente de su productividad y los trazó contra el número acumulativo de artículos. Esto dió como resultado una curva que se presenta en el cuadro 1.

CUADRO 1



La curva sube gradualmente y después se convierte en una línea recta. La posición x, en donde la curva se convierte en una línea recta, es el límite superior del núcleo, en donde se incluyen las revistas más productivas en relación con artículos publicados en una determinada materia.

INFOBILA

Bradford definió el núcleo como (1) aquellas revistas que producen mas que 4 artículos al año; y además distinguió otros dos grupos como (2) aquellas revistas que tenian más de un artículo pero menos que cuatro en el año y (3) aquellas revistas que producian un artículo al año. Las delimitaciones grupales las hizo en tal manera que cada grupo tuviera, aproximadamente, el mismo número de artículos. Bradford observó también que los números de revistas en los tres grupos, formaban una progresión geométrica de la forma $1:n:n^2$. Esto significa que el número de revistas en un grupo es n veces el número de revistas del grupo anterior.

MODIFICACIONES DE LA LEY DE BRADFORD

La ley de Bradford ha sido modificada por varios investigadores entre los cuales mencionaremos, en forma breve, a Vickery, Cole, Leimkuhler y Groose.

Vickery reconoció que Bradford en su apuro por enfatizar la parte lineal de la curva, presentó una formulación matemática que violaba su propia lógica. Por lo tanto, Vickery corregió la formulación de Bradford en tal forma que tomó en consideración ambos aspectos del modelo: la parte lineal y la parte curva.

El argumento de Vickery para esta modificación es de la siguiente manera: si $R(n)$ es el acumulativo total de referencias (citas) relevantes encontradas en las primeras n revistas ordenadas 1, 2, 3,... en orden decreciente de productividad, la formulación de Bradford debe ser:

$$R(n) = R(n^2) - R(n) = R(n^3) - R(n^2) = R(n^4) - R(n^3) = \ldots \quad \text{Para } n \gg 1$$

Además argumentó que con esta formulación la Ley de Bradford puede aplicarse a cualquier número de zonas.

Cole, después de ordenar las revistas en orden decreciente de citas, graficó el logaritmo de las revistas más productivas (x) contra la fracción $F(x)$ de las citas encontradas en esas revistas y expresó esto en la siguiente fórmula que es muy similar a una ecuación lineal:

$$F(x) = 1 + K \log x$$

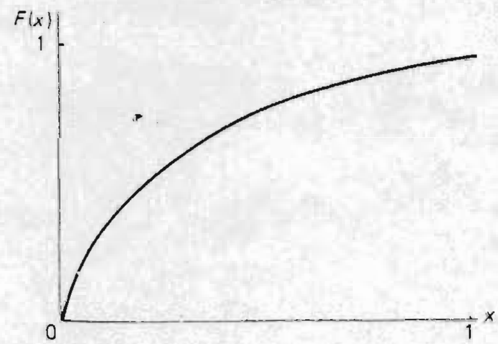
Además sugirió que la pendiente K es una medida precisa y útil del grado de dispersión de las citas, la cual podría ser característica de cada campo temático.

Leimkuhler, mostró que $F(x)$, de la ecuación anterior, es una aproximación de la Ley de Bradford y que la forma exacta de la función de distribución de $F(x)$ de citas, en cada fracción de las revistas mas productivas, debe ser:

$$F(x) = \frac{\log_e (1+Bx)}{\log_e (1+B)} = \frac{\log_{10} (1+Bx)}{\log_{10} (1+B)}$$

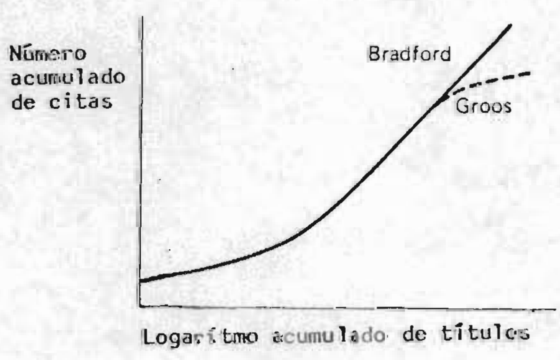
donde B es una constante relacionada tanto con la amplitud de la colección, como con el tema. Del cuadro 2, se puede observar que una pequeña fracción de las revistas más productivas corresponde a una proporción alta de éstas.

CUADRO 2



Groos, analizó y graficó más de 20,000 artículos de revistas en física, según el formato de Bradford, y observó que en algún punto en la suma acumulativa de títulos se producía una curva hacia abajo que se alejaba de la ley lineal de Bradford. Brookes ha nombrado a este fenómeno la "caída de Groos" y ha dado su explicación al respecto: Si se encuentran nuevas revistas que contienen artículos relevantes la "caida de Groos" disminuye. Por lo tanto, la caida puede reflejar el resultado de una búsqueda incompleta.

CUADRO 3



Brookes manipuló matemáticamente la formulación de Vickery y presentó las siguientes ecuaciones:

$$R(n) = \alpha n^{\beta} \quad \text{para} \quad 1 \leq n \leq c$$
$$R(n) = N \log_e (n/s) \quad \text{para} \quad c \leq n \leq N$$

en donde:

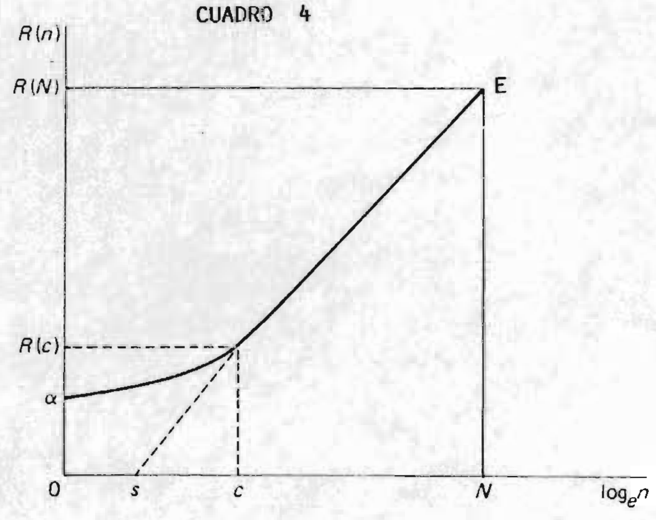
$R(n)$ = suma acumulativa de citas en n revistas
α = el número de citas en la revista más productiva.
β = una constante menor de 1
N = el número total de revistas que contienen artículos relevantes (estimado)
S = intersección de la línea recta de Bradford con el eje de las x y su valor es mayor que 1
C = el punto sobre el eje de las x donde empieza la porción linear de la curva.

El número total (estimado) de artículos, sin tomar en cuenta la cantidad de artículos que se haya encontrado, está dado por:

$$R(N) = N \log_e (N/S)$$

Estos resultados permitirán estimar, aún con una búsqueda incompleta, el número total de revistas que contienen artículos citados y el número total de artículos. Esto quiere decir, que si se colectan datos suficientes para lograr la línea recta, puede predecirse el fin de la misma (E en el Cuadro 4].

CUADRO 4



UN CASO PRACTICO DE APLICACION DE BRADFORD

Para demostrar la funcionalidad de la Ley de Bradford se presenta en seguida un estudio que se llevó a cabo en la Biblioteca Daniel Cosío Villegas.

Para tal efecto, se analizaron las formas que los lectores tenían que llenar para solicitar en préstamo las revistas. En el recuento de título se excluyeron los censos y los informes anuales.

Durante el período de estudio se contabilizaron 1763 consultas distribuidas entre 267 títulos. El cuadro 5 muestra el ordenamiento de los resultados obtenidos. Como se puede observar había un título que tenía 159 circulaciones en contraste con 93 títulos que circularon una vez en el período. La columna C es el producto del número de títulos en cada categoría por su circulación correspondiente.

La columna D es el acumulativo de las circulaciones obtenidas hasta ese renglón, la columna E es el acumulativo de títulos hasta ese renglón y la columna F es el logaritmo natural del número correspondiente en la columna E.

Las zonas a las que Bradford hace mención, se construyen en un proceso de intento y error para dividir el total de circulaciones en grupos que contengan, aproximadamente, el mismo número de circulaciones. Según Goffman y Morris, el número más pequeño para intentar la zonificación, debe tener más de la mitad del número de revistas quecircularon una sola vez (en este caso, aquel número es 47). Sin embargo, como el primer título circuló 159 veces y no se pueden considerar fracciones de título, la zonificación más pequeña tendría que ser de tamaño 159, al tiempo que la siguiente opción sería de 264, que son las circulaciones de los dos primeros títulos.

El objetivo de la determinación de zonas es el de encontrar n, conocido como multiplicador de Bradford, y que es el factor importante en el enunciado de la ley de este autor. El proceso de búsqueda de n, también permite determinar el núcleo. En la práctica el cuadro 5, con sus dos primeras columnas, sería suficiente para ayudar a la toma de decisiones sobre: qué comprar, qué encuadernar, qué descartar, de qué títulos necesitaría un segundo ejemplar, etc.

Una ilustración del proceso iterativo para determinar el tamaño de las zonas, puede consultarse en Quijano y Villarreal. Para efectos de este trabajo, se presentan los resultados del caso de la Biblioteca Daniel Cosío Villegas en el cuadro 6, mientras que en el cuadro 7, se ilustra la gráfica obtenida con esos resultados.

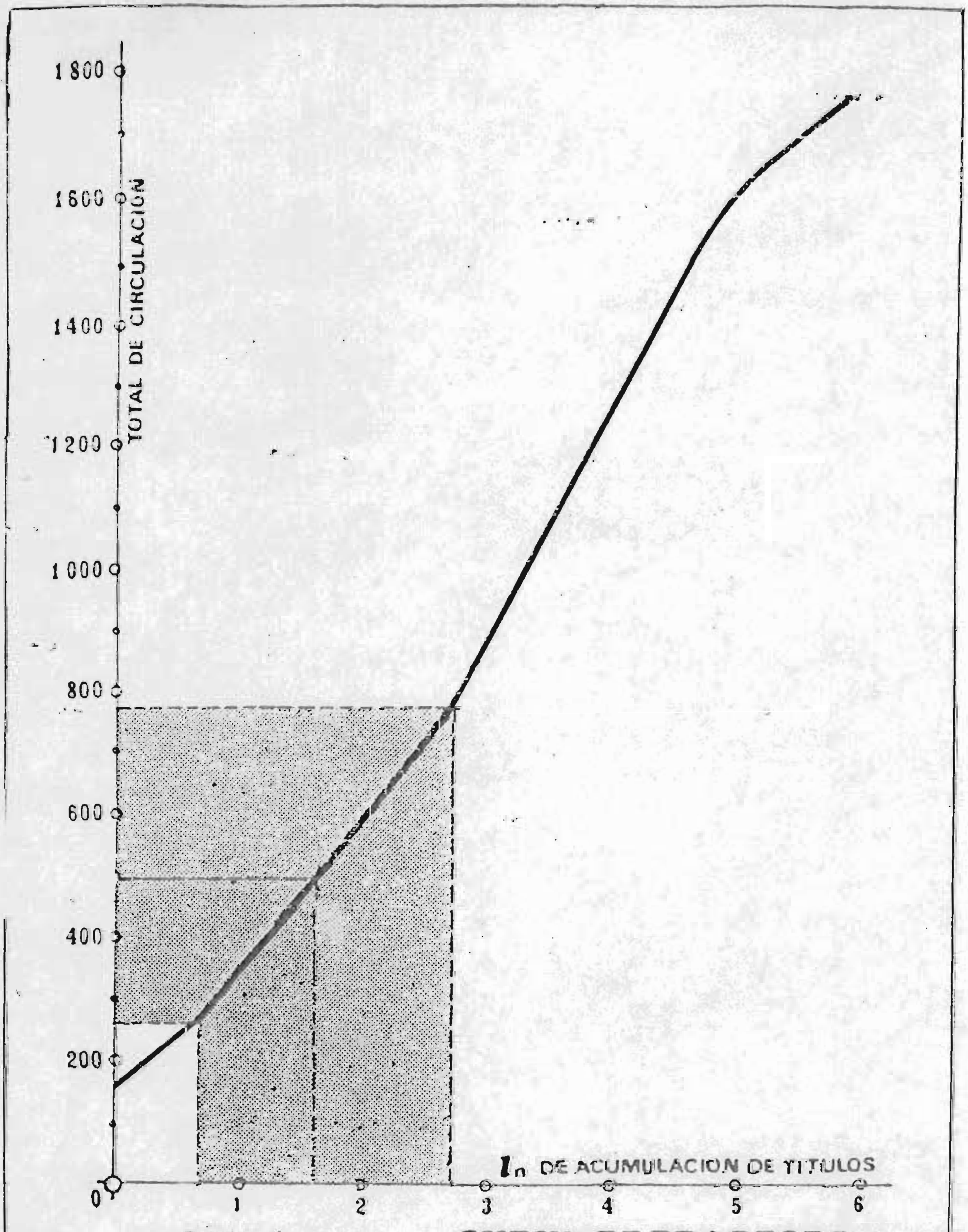
A Número de Títulos	B Número de Circulaciones en Cada Título	C Total de Circulaciones en Cada Categoría (A x B)	D Número Acumulado de Circulaciones	E Número Acumulado de Títulos	F Logaritmo de E
1	159	159	159	1	0
1	105	105	264	2	0.69
1	96	96	360	3	1.1
1	77	77	437	4	1.39
1	53	53	490	5	1.61
2	48	96	586	7	1.95
1	28	28	614	8	2.08
1	27	27	641	9	2.2
1	25	25	666	10	2.3
1	22	22	688	11	2.4
1	21	21	709	12	2.48
3	20	60	769	15	2.71
3	19	57	826	18	2.9
2	18	36	862	20	3.0
4	15	60	922	24	3.18
2	14	28	950	26	3.26
2	13	26	976	28	3.33
4	12	48	1024	32	3.46
8	10	80	1104	40	3.69
3	9	27	1131	43	3.76
9	8	72	1203	52	3.95
11	7	77	1280	63	4.14
14	6	84	1364	77	4.34
15	5	75	1439	92	4.53
21	4	84	1523	113	4.73
25	3	75	1598	138	4.93
36	2	72	1670	174	5.16
93	1	93	1763	267	5.59

Cuadro 5

Zona	Número de consultas en la zona	Número de títulos en la zona	Multiplicador de Bradford (n)	Prueba
1	264	2	---	---
2	274	4	2	4.1
3	250	10	2.5	8.4
4	266	19	1.9	17.23
5	256	33	1.7	35.32
6	249	57	1.7	72.4
7	204	142	2.5	148.44
	1,763	267	12.3	

(n) promedio = 2.05

Cuadro 6 Distribución de Bradford para los títulos consultados.



Cuadro 7 **CURVA DE BRADFORD**

... para la muestra estudiada

MODELOS MORSE-MARKOV

El primer intento de utilizar los modelos probabilisticos para la predicción del uso de libros fue realizado por Philip Morse del Massachussetts Institute of Technology en 1962. Morse aplico la cadena de Markov, como base, para desarrollar sus modelos de evaluación de obras monográficas que posteriormente fue nombrado como los modelos matemáticos Morse-Markov. Para dar una idea general y breve sobre la cadena de Markov se da su definición junto con un ejemplo en los siguientes párrafos.

Una cadena de Markov es un proceso aleatorio que tiene la probabilidad especial de que se puede predecir su futuro con la misma exactitud a partir del conocimiento de la situación actual que a partir del conocimiento del presente junto a todo el pasado. Considérese el siguiente ejemplo para ver la funcionalidad de este concepto en una situación real.

Supongamos que una empresa de mercadotecnia quiere desarrollar un modelo probabilistico sobre 3 productos y observar su comportamiento relativo a su venta en el mercado. Para simplificar se numeran a los tres productos como 1, 2 y 3 y el acto de compra de estos tres productos como a1, a2 y a3. Además, se establece la probabilidad de compra de estos tres productos en base a ventas anteriores y estudios de mercado. Estas probabilidades reflejan la compra de estos productos por un consumidor ordinario. Se ha encontrado que, por lo general, un consumidor tiene la tendencia de comprar el mismo producto que habia comprado en el pasado. Aún más, se supone que la compra de un producto es condicionado, solamente, en la compra del último producto y no en las compras anteriores.

En el cuadro 8 se presentan las probabilidades de los tres productos.

Ultima compra	Proxima compra a1	a2	a3
a1	0.60	0.20	0.20
a2	0.10	0.70	0.20
a3	0.30	0.20	0.50

La probabilidad que la próxima compra sea el producto 1 dado que la última compra era también el producto 1 se anota como P(a1/a1)=0.60; la probabilidad de cambio del producto 1 al producto 2 se indica como P(a2/a1)=0.20; y la probabilidad de cambio del producto 1 al producto 3 se muestra como P(a3/a1)=0.20. Asi se puede indicar la compra de los tres productos, dado que las compras anteriores eran los productos 2 o 3. Este proceso de Markov se llama también la probabilidad transicional y se explica su aplicación en relación al uso del libro un poco más adelante.

Morse define tres supuestos relativos a la circulación de libros:

1) El proceso de la circulación de libros es aleatorio
2) La circulación de libros decrece con el tiempo
3) Hay una correlación de tiempo de un período al próximo

Los dos primeros supuestos han sido abordados, suficientemente, en la literatura y sus explicciones son redundantes en este trabajo. Sin embargo, se da una descripción breve sobre el supuesto 3.

El simple modelo probabilístico que exhibe una correlación de tiempo se conoce como el proceso de Markov. En este proceso el estado del sistema al final de un período dado, se determina por su estado al principio del periodo. Por ejemplo, si un grupo de libros obedece el proceso de Markov, entonces la probabilidad de que un libro que circuló m veces en el año t circulará n veces en el año t+1 es Tmn conocido como la probabilidad transicional. Esto significa el tránsito de m circulación de un libro en el año t a n circulación en el año t+1. El valor de Tmn depende de m circulación en el año t pero indirectamente de la historia de circulación en los años anteriores.

Para obtener la matriz de probabilidad transicional de un grupo de libros es necesario reconstruir la historia de la circulación de éste y realizar algunas manipulaciones matemáticas. Obsérvese el cuadro 9 de la clase 300-199 de la biblioteca del Instituto Tecnológico Autónomo de México (ITAM). Este cuadro se construyó por los datos de las tarjetas de circulación y un poco por manipulación numérica.

m	M(m)	N mn n=0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	M(m)	Teor.
0	353	156	78	43	43	16	2	8	4	0	1	0	0	1	0	0	0	1	1.36	2.01
1	242	56	62	39	30	13	9	7	8	3	5	1	3	0	0	0	1	0	2.36	2.23
2	117	28	23	17	15	6	12	6	3	3	0	1	0	1	1	1	0	0	2.70	2.46
3	98	19	8	18	13	6	10	6	4	4	5	1	1	2	0	1	0	0	3.64	2.68
4	55	6	8	11	15	3	1	4	2	3	1	1	0	0	0	0	0	0	3.15	2.91
5	46	9	5	2	5	7	7	4	2	0	0	2	2	0	0	1	0	0	3.93	3.13
6	25	6	3	2	3	3	0	3	4	0	0	0	0	1	0	0	0	0	3.44	3.35
7	20	4	3	2	1	0	2	3	1	0	1	1	1	1	0	0	0	0	4.35	3.58
8	12	4	0	2	1	0	2	0	1	1	0	0	1	0	0	0	0	0	3.58	3.80
9	6	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	3.50	4.02
10	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	4.25
11	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	5.00	4.47
12	2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	4.00	4.70
13	2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2.50	4.92
14	2	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	6.50	5.14
15	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8.00	5.37

ALFA = 2.011 BETA = 0.224

Los símbolos de este cuadro se definen de la siguiente manera:

- m= número de la circulación en el año t
- n= número de la circulación en el año t+1
- M(m)= número de libros con m circulación en el año t
- Nmn= número de libros con m circulación en el año t y n circulación en el año t+1
- N(m)= el promedio de la circulación en el año t+1 dado m circulación en el año t

Si trazamos los valores de N(m) contra los valores de m, se obtiene una relación líneal que se expresa en la forma:

$$N(m)=\alpha+\beta m$$

Se puede observar que esta ecuación es la de la línea recta, es decir Y=a+bx, siendo α y β intersección y pendiente respectivamente. El cálculo de α y β se obtiene por la técnica de mínimos cuadrados con las siguientes fórmulas

$$\begin{cases} a=\dfrac{\Sigma Y-b\Sigma X}{n} \\ \\ b=\dfrac{n\Sigma XY-\Sigma X\Sigma Y}{n\Sigma X^2-(\Sigma X)^2} \end{cases}$$

Aplicando las ecuaciones anteriores y utilizando los valores de m y N(m) de la clase 300-399 se dan los siguientes resultados:

m (X)	N(m) (Y)	m.N(m) (XY)	m^2 (X^2)
0	1.36	0.00	0
1	2.36	2.36	1
2	2.70	5.40	4
3	3.64	10.92	9
4	3.15	12.60	16
5	3.93	19.65	25
6	3.44	20.64	36
7	4.35	30.45	49
8	3.58	28.64	64
9	3.50	31.50	81
10	1.00	10.00	100
11	5.00	55.00	121
12	4.00	48.00	144
13	2.50	32.50	169
14	6.50	91.00	196
15	8.00	120.00	225

$\Sigma X=120$ $\Sigma Y=59.01$ $\Sigma XY=518.66$ $\Sigma X^2=1240$ $(\Sigma X)^2=14400$ $n=16$

$$\beta=\frac{(16\times518.66)-(120\times59.01)}{(16\times1240)-(14400)}=0.224$$

$$\alpha=\frac{(59.01)-(0.224\times120)}{16}=2.001$$

INFOBILA

Se mencionó anteriormente que Morse definió que el proceso de la circulación de libros era aleatorio. Por lo tanto, la distribución de circulación puede tomarse como la distribución de Poisson:

$$Tmn = \frac{(\alpha+\beta m)^{n}}{n!} e^{-(\alpha+\beta m)}$$

Con los valores diferentes de m y n en esta fórmula se da la siguiente matriz de probabilidad transicional de la clase 300-399 para el año t=1.

Matriz de Probabilidad Transicional de la clase 300-399, t=1

BETA = .224 ; ALFA = 2.011

Tmn

m \ n =	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0.134	0.269	0.271	0.181	0.091	0.037	0.012	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.107	0.239	0.267	0.199	0.111	0.050	0.019	0.006	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.086	0.210	0.259	0.212	0.130	0.064	0.026	0.009	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.068	0.183	0.246	0.220	0.148	0.079	0.035	0.014	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.055	0.159	0.231	0.224	0.163	0.095	0.046	0.019	0.007	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000
5	0.044	0.137	0.214	0.223	0.175	0.110	0.057	0.026	0.010	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000
6	0.035	0.117	0.196	0.220	0.184	0.124	0.069	0.033	0.014	0.005	0.002	0.001	0.000	0.000	0.000	0.000	0.000
7	0.028	0.100	0.179	0.213	0.191	0.137	0.081	0.042	0.019	0.007	0.003	0.001	0.000	0.000	0.000	0.000	0.000
8	0.022	0.085	0.161	0.204	0.194	0.148	0.094	0.051	0.024	0.010	0.004	0.001	0.000	0.000	0.000	0.000	0.000
9	0.018	0.072	0.145	0.194	0.195	0.157	0.106	0.061	0.031	0.014	0.006	0.002	0.001	0.000	0.000	0.000	0.000
10	0.014	0.061	0.129	0.182	0.194	0.165	0.117	0.071	0.038	0.018	0.008	0.003	0.001	0.000	0.000	0.000	0.000
11	0.011	0.051	0.114	0.170	0.190	0.170	0.127	0.081	0.045	0.023	0.010	0.004	0.002	0.001	0.000	0.000	0.000
12	0.009	0.043	0.101	0.157	0.185	0.174	0.136	0.091	0.054	0.028	0.013	0.006	0.002	0.001	0.000	0.000	0.000
13	0.007	0.036	0.088	0.145	0.178	0.175	0.144	0.101	0.062	0.034	0.017	0.008	0.003	0.001	0.000	0.000	0.000
14	0.006	0.030	0.077	0.132	0.170	0.175	0.150	0.110	0.071	0.041	0.021	0.010	0.004	0.002	0.001	0.000	0.000
15	0.005	0.025	0.067	0.120	0.161	0.173	0.155	0.119	0.080	0.048	0.026	0.012	0.006	0.002	0.001	0.000	0.000
16	0.004	0.021	0.058	0.108	0.152	0.170	0.158	0.127	0.089	0.055	0.031	0.016	0.007	0.003	0.001	0.000	0.000

EXTENSION DE LOS MODELOS TEORICOS DE MORSE

Como se mencionó con anterioridad Morse utilizó el concepto de las cadenas de Markov en la circulación de libros argumentando que la circulación en el año t+1 depende de la circulación en el año t. El nombró los números de circulación como n y m respectivamente. Aceptando este supuesto, era por lo tanto, necesario tener los datos correspondientes a la circulación de un grupo de libros en el primer año. Por ende, los modelos de Morse-Markov fueron modificados de tal manera que una muestra pequeña de una clase de libros durante un tiempo definido, permitiría predecir el comportamiento de ésta sin considerar la circulación en el primer año de la vida del libro.

Un ejemplo aclarará este aspecto. Supongamos que se toma una muestra de una clase de libro para observar su comportamiento y predecir la circulación de éste en los años posteriores. Los modelos Morse-Markov permitirían hacer esto siempre y cuando la historia de la circulación de este muestreo esté disponible durante el primer año después de su adquisición. En la extensión de los modelos Morse-Markov, se registra la circulación de una clase de libro por uno o dos meses y se calculan probabilisticamente la circulación de los meses faltantes para obtener la circulación de un año. Por ejmplo, en el cuadro 10 el valor de M(J)= 372 son libros con la circulación J=1,2,... de un mes. El valor de N9J)= 3189 son libros con las circulaciones correspondientes de un año. Además la extensión del modelo permitiría calcular el valor de N(J,O)= 4152 que son libros activos y más libros que tienen potencial para circular.

DISTRIBUCION NO CORREGIDA DE LA CLASE 300-399

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	suma
M(j)	77	76	73	32	27	27	19	11	9	6	8	4	0	1	1	1	372
N(j)	1407	727	488	168	118	103	65	34	26	16	20	10	0	2	2	2	3189
N(j,0)	2226	841	513	171	119	103	65	34	26	16	20	10	0	2	2	2	4152

Las fórmulas utilizadas en la extensión de modelos Morse-Markov son:

$Na = \sum N(J)$ libros activos que circularon en tiempo de muestreo

$NM = \sum N(J,O)$ libros activos más libros que tienen potencial para circular

$NP = NM - Na$ número de libros potenciales

$\sum JN(J) =$ circulación de libros a diario

$Ra = \frac{Na}{N(1)}$ el promedio de circulación de libros a diario

$RaNa = R_M M$ circulación anual

Aplicando la probabilidad transicional y los modelos correspondientes se puede observar el comportamiento de un grupo de libros en los años posteriores. El cuadro 11 muestra el comportamiento de una clase de libros en los años 0,1,2,3 de acuerdo con modelos Morse-Markov.

Los parámetros y valores de la clase 300-399 en los años* (0), (1), (2) y (3)

Año	α_a	α_M	β	γ	N_a	N_M	$C=N_a/N_M$	R_a	P_M=	$R_M N_M$
(0)	2.011	1.544	0.224	0.559	3189	4152	0.768	2.27	1.74	7224
(1)	2.011	1.544	0.224	0.563	3499	4152	0.846	2.29	1.93	8013
(2)	2.011	1.544	0.224	0.569	3544	4152	0.852	2.32	1.98	8221
(3)	2.011	1.544	0.224	0.571	3546	4152	0.853	2.33	1.99	8262

* año (0)=Octubre 1, 1983-Septiembre 30, 1984

año (1)=Octubre 1, 1984-Septiembre 30, 1985

año (2)=Octubre 1, 1985-Septiembre 30, 1986

año (3)=Octubre 1, 1986-Septiembre 30, 1987

CONCLUSIONES

1. La teoría del desarrollo de colecciones sobre la base de evaluar los materiales bibliográficos por su demanda, su valor y su uso, encuentra en los modelos presentados una herramienta auxiliar para la toma de decisiones y el establecimiento de políticas. Es importante señalar, sin embargo, que estas herramientas deberán estar siempre supeditadas a l intuición administrativa y a la experiencia como ponderadora de factores cualitativos.

2. En su momento, ambos modelos fueron aplicados a realidades concretas y fueron propuestos y utilizados como instrumentos de decisión.

3. A pesar de la complejidad de los planteamientos matemáticos, tanto la distribución de Bradford como las cadenas de Markov son fáciles de interpretar, toda vez que expresan principios de sentido común para el bibliotecario experimentado.

4. En contraste con el momento en que ambos modelos fueron utilizados, la reciente tendencia a la automatización de procesos y servicios bibliotecarios favorece el cálculo y el desarrollo analítico de las fórmulas. Es importante que las bibliotecas consideren el levantamiento de este tipo de datos como parte de sus sistemas de información administrativa.

5. Ambos modelos conllevan la posibilidad de tomar decisiones en cuanto a la racionalización de las presiones que se ejercen sobre la biblioteca para adquirir materiales de manera indiscriminada. En otras palabras, los resultados de estos ejercicios permitirán jerarquizar la asignación de recursos.

6. Un problema que podría ser resuelto con este tipo de instrumentos es el de ubicar acervos, sin descartarlos, en lugares remotos en donde el almacenamiento resulte mas barato.

7. Los bibliotecarios deberían incorporar a su acervo profesional la interpretación de los resultados que emanan de los estudios métricos, como una manera de tomar decisiones sobre bases científicas.