

As armadilhas da lei de Bradford

Eliana Coutinho

Rede Ferroviária Federal S.A.

Gerência de Documentação

20224 Rio de Janeiro, RJ

Resumo – Trata das diferenças entre o modelo teórico e os estudos empíricos da lei de Bradford e as consequentes distorções dos resultados destes últimos, quando a lei é mal utilizada.

1 Introdução

O termo “Bibliometria” foi usado pela primeira vez com o sentido de “aplicação de matemática e métodos estatísticos a livros e outros meios de comunicação” por Pritchard(1) em 1969, quando reuniu as três leis clássicas da Bibliometria:

- 1) Lei de Bradford, originária de 1934, sobre produtividade de periódicos em relação a determinado assunto;
- 2) Lei de Zipf, de 1929, sobre freqüência do uso de palavras em um texto; e
- 3) Lei de Lotka, de 1934, sobre produtividade de autores.

A lei de Bradford é a mais conhecida e a mais usada em estudos bibliométricos, como auxílio na administração de sistemas de informação bibliográfica.

O fenômeno objeto da lei, ou seja, o da distribuição de artigos sobre determinado assunto em periódicos científicos, é fácil de ser observado. No entanto, dependendo da situação analisada, a aplicação da lei, sem maiores cuidados, pode levar a resultados pouco confiáveis, já que o modelo atualmente disponível ainda é insuficiente para identificar o verdadeiro mecanismo do fenômeno.

Como a Bibliometria é uma área recente, muitos estudos ainda serão necessários para que seus métodos sejam bem estruturados, os fenômenos observados no comportamento da literatura científica e técnica sejam bem entendidos, e assim possa ser usada como um instrumento eficaz nos vários serviços dos sistemas de informação bibliográfica.

2 Lei de Bradford

Bradford(2), especialista da informação da Biblioteca do Museu de Ciências, na Inglaterra, apresenta observações a respeito da dispersão de artigos sobre assunto específico em vários periódicos devotados a assuntos diferentes.

Em 1948, Bradford reuniu no livro(3) uma coleção de ensaios sobre a arte da documentação. Segundo prefácio do próprio autor(4), tratava-se do primeiro trabalho em língua inglesa sobre o assunto e o terceiro publicado a esse respeito no mundo, precedido pelo **Traité de Documentation: le livre sur le livre**, de Paul Otlet, publicado em 1934, em Bruxelas, e pelo **Die De-cimal Klassifikation**, de O. Frank, publicado em Berlim em 1946.

No capítulo 'O caos documentário'(5), que é a reprodução do artigo publicado no periódico *Engineering*(2), já mencionado, Bradford chama a atenção para o caos bibliográfico resultante da desorganização existente na indexação por assunto e nos serviços de resumos, pois muitos documentos relevantes deixam de ser incluídos em periódicos de resumos e em bibliografias.

Através de sua experiência com periódicos da Biblioteca de Ciências, Bradford verificou que os periódicos de resumos divulgavam menos da metade dos documentos úteis publicados. Logo, mais da metade dos documentos ficava sem utilização pelos usuários.

Segundo o autor, essas omissões seriam decorrentes do fato dos serviços de resumos se concentrarem nos periódicos dedicados ao assunto de seu interesse específico, acrescentando somente alguns poucos especializados em outras áreas ligadas mais de perto a esse assunto. Isto não seria suficiente para extrair a maioria dos artigos publicados sobre o assunto. Além disto, livros, folhetos, especificações de patentes e semelhantes, eram ignorados pelos serviços de resumos.

Bradford sugere a necessidade de se identificar a dispersão dos artigos sobre determinado assunto publicados em periódicos de outras especialidades, já que as áreas científicas se interrelacionam, para que maior número de artigos relevantes possam ser recuperados, além daqueles publicados em periódicos dedicados especificamente à área de interesse.

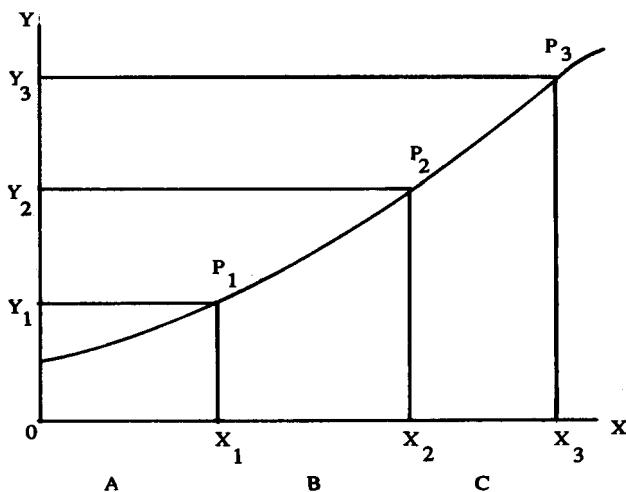
A dispersão é verificada a partir da identificação do núcleo de periódicos dedicados a determinado assunto, que é formado por poucos títulos produtores de grande número de artigos de interesse. Este núcleo é a zona de produtividade nº 1 da distribuição. Os outros periódicos menos produtivos são ordenados em zonas de produtividade decrescente de artigos sobre o mesmo assunto, e o número de periódicos em cada zona aumenta à medida que sua produtividade individual diminui. O número de artigos em cada zona, no entanto, permanece mais ou menos constante.

A lei de distribuição de artigos sobre determinado assunto em periódicos científicos foi assim enunciada por Bradford(5):

"Se os periódicos científicos forem ordenados em ordem de produtividade decrescente de artigos sobre determinado assunto, poderão ser distribuídos em um núcleo de periódicos mais particularmente dedicados a esse assunto e em diversos grupos ou zonas contendo o mesmo número de artigos que o núcleo, sempre que o número de periódicos do núcleo e das zonas sucessivas for igual a $1 : n : n^2 \dots$ "

A lei de Bradford traz embutida uma conclusão comum a outras leis da Bibliometria: poucos produzem muito e muitos produzem pouco.

Bradford também elaborou um gráfico semi-logarítmico para ilustrar sua lei, tendo representado no eixo da abcissa o número dos periódicos em ordem decrescente de produtividade de artigos relevantes a determinado assunto (x) e no eixo da ordenada, o número total acumulado de artigos (y), marcando três zonas de produtividade (P_1, P_2, P_3), sendo P_1 o núcleo:



Para Bradford, os periódicos estariam classificados em três classes representadas no gráfico anterior, que seriam:

- A – os que produzem mais de quatro referências por ano;
- B – os que produzem mais de uma e não menos de quatro por ano;
- C – os que produzem uma ou menos por ano.

Mais tarde, Vickery(6) menciona que a distribuição dos periódicos sobre determinado assunto poderia ter qualquer número de zonas, e não apenas três, conforme indicado por Bradford.

Goffman & Warren(7) alteraram a distribuição esticando o número de zonas para o máximo possível, após pesquisa com base na análise matemática de literatura médica.

3 Desenvolvimento teórico e prático da lei de Bradford

A lei de Bradford é uma lei empírica e foi, por cerca de 20 anos, considerada como uma curiosidade estatística, porque não era expressa sob a forma de equação matemática, ou seja, não tinha base científica.

A formulação de uma lei, sob ponto de vista científico, permite prever acontecimentos relacionados ao fenômeno que descreve e ajuda no desenvolvimento de teorias que expliquem as razões da ocorrência de determinados padrões.

O desenvolvimento teórico da lei de Bradford baseia-se na explicação dos eventos probabilísticos que se juntam para criar o padrão regular de dispersão de artigos em vários periódicos. Para isto é necessário, em primeiro lugar, descrever esse padrão matematicamente.

A primeira fórmula foi dada pelo próprio Bradford: $1 : n : n^2 \dots$. No entanto, por se tratar de lei empírica, não era derivada estritamente de dados, apresentando várias falhas, como o fato da dispersão ser expressa em termos de grupos de periódicos (zonas de produtividade), não em termos de produtividade individual dos títulos; e a falha de explicação das causas da dispersão.

Vinte anos após sua formulação, a lei de Bradford teve sua base teórica desenvolvida por Brookes(8), que a partir do trabalho de Vickery(6), elaborou uma fórmula para calcular o número de artigos produzidos por um periódico, conforme o lugar que ocupa na ordenação por uso:

$R(n) = K \log (n)$,

Onde:

n é o lugar que o periódico ocupa na ordem de série por uso.

O valor de R(1) é o número de artigos do periódico mais usado, R(2) é a soma do número de artigos produzidos pelo primeiro, mais os do segundo periódico mais usado, e assim por diante. K é uma constante, conforme a coleção examinada.

Com essa fórmula acusava falha na previsão, isto é, determinava número alto demais de artigos para os periódicos do núcleo (os mais freqüentes) e baixo demais para o outros, Brookes modificou-a para:

$R(n) = K \log (n/s)$,

Onde s é uma constante que representa a medida de amplitude do assunto objeto da pesquisa.

Essa modificação se revelou inadequada para os periódicos mais produtivos, o que sugere que os processos que criam a regularidade de comportamento destes periódicos podem ser diferentes dos que determinam o comportamento dos pouco produtivos.

A fórmula matemática é aplicada para periódicos que contribuem com poucos artigos. Para os teóricos, eles são importantes porque demonstram o desenvolvimento da área à medida que atraem mais artigos a respeito dela, de acordo com o mecanismo de "sucesso traz sucesso", tratado por Price(9) no seu modelo de vantagens acumuladas. Esses periódicos são considerados a base sobre a qual se constrói a distribuição de Bradford.

Os estudos empíricos da lei de Bradford têm como objetivo buscar uma forma de usar a regularidade quantitativa para tomada de decisão, com bases mais precisas do que a simples intuição. Esses estudos permitem identificar as implicações do padrão observado melhoramento dos serviços de atendimento ao usuário.

Para Drott(10) os estudos empíricos envolvem dois aspectos:

1) previsão: que títulos serão úteis, ou como os usuários podem comportar-se em relação ao uso da coleção;

2) avaliação: medição dos serviços de recuperação ou aquisição através dos padrões teóricos.

Assim, a lei de Bradford é usada, na prática, para planejamento e administração de sistemas de informação. Por isso, sob o ponto de vista empírico, o mais importante é identificar os periódicos mais produtivos ou mais

usados, para estabelecimento de política de aquisição racional e recuperação de informações relevantes.

4 As armadilhas da lei de Bradford

Comparando os dois pontos de vista – o teórico e o empírico, nota-se o que Drott(10) chama de “lacuna” entre os estudos teóricos e empíricos da lei de Bradford.

Os teóricos acreditam que o avanço do entendimento depende do estudo de certas fórmulas matemáticas e que sua conformidade com dados empíricos é menos importante.

O modelo teórico apresentado para a lei não se ajusta aos estudos empíricos porque não prevê as variáveis existentes numa situação real analisada. Tais variáveis são referentes ao assunto estudado, às formas pelas quais o estudo é realizado, às necessidades específicas do usuário ou às características da coleção.

Os números totais de artigos e de periódicos usados no estudo empírico estão relacionados respectivamente com as constantes K e s da fórmula no modelo teórico. Qualquer variável na situação empírica que afete esses totais causará um desajuste ao modelo teórico, podendo levar o estudo a conclusões inadequadas.

Essa lacuna entre os estudos teóricos e empíricos é o elemento responsável pelas armadilhas que a lei de Bradford pode provocar na sua aplicação ao estudo do comportamento da literatura sobre determinado assunto.

O fato do modelo teórico não permitir que a situação empírica seja efetivamente representada exige que a lei seja usada com muito cuidado, pois os resultados e estudos mal realizados podem levar a ações administrativas incorretas por parte dos gerentes de sistemas de informações. Decisões referentes a seleção, aquisição e descarte de coleções de periódicos baseadas em conclusões distorcidas pelo uso descuidado da lei de Bradford se refletem diretamente na queda de qualidade do atendimento ao usuário.

A recuperação de informações relevantes é a própria razão de ser de qualquer sistema de informação, e a lei de Bradford é um instrumento aparentemente eficaz e simples de ser usado para esse fim. No entanto, é preciso muito cuidado ao aplicá-la, já que estudos mal estruturados podem comprometer seriamente um sistema de informação.

Para evitar as armadilhas que a lei de Bradford pode apresentar nos estudos empíricos, o pesquisador deve estar atento aos elementos que caracterizam a situação real a ser estudada. Esses elementos são as variáveis a serem consideradas na aplicação clássica da lei, como: generalidade ou especificidade da área, política editorial dos periódicos, periodicidade dos títulos examinados, período de tempo estudado etc.

Enquanto não houver estudo teórico suficiente sobre a lei, a fim de que seja possível incorporar variáveis empíricas ao modelo teórico, existirá o risco de se cair nas armadilhas de uma pesquisa mal elaborada.

Pinheiro(11) percebeu uma dessas armadilhas da lei de Bradford e testou a seguinte hipótese em sua dissertação de Mestrado:

“o núcleo de periódicos determinado pela lei de Bradford não corresponde ao dos ‘mais particularmente devotados ao assunto’, e sim aos periódicos mais produtivos em determinado período de tempo.”

O assunto escolhido para estudo foi “classificação”.

A partir de artigos coletados no Library and Information Science Abstracts (LISA), no período de 1974 e 1978, foi feita a distribuição clássica de Bradford, que demonstrou a produtividade absoluta dos periódicos, isto é, a totalidade de artigos sobre classificação publicados em cada periódico, obtendo-se, após, a ordenação em série dos títulos dos periódicos por produtividade absoluta decrescente.

Para testar a hipótese da dissertação foi preciso verificar a produtividade relativa dos periódicos, ou seja, a “proporção entre o número de artigos sobre Classificação (produtividade absoluta) e o total de artigos publicados pelo periódico no período, em todos os campos”.

Calculada a produtividade relativa, os títulos dos periódicos foram reordenados de acordo com a porcentagem obtida pelo número de artigos sobre Classificação em relação ao número total de artigos publicados em cada periódico.

Observou-se alteração na ordenação original, obtida através da distribuição clássica de Bradford, de quase todos os títulos, confirmado-se a hipótese de que os periódicos determinados pela lei como os mais devotados ou especializados no assunto objeto da pesquisa, na verdade representam os mais produtivos, porque a lei não considera variáveis como: periodicidade, número de páginas e, principalmente, o percentual correspondente ao grau de devotamento ou especificação de cada periódico a um determinado assunto.

Com esta conclusão, Pinheiro(11) lança uma reformulação conceitual da lei de Bradford, que se reflete diretamente na administração de sistemas de informação, além de representar uma grande contribuição para os estudos teóricos e empíricos sobre o assunto.

5 Conclusão

Em relação à lei de Bradford, existem duas linhas de estudo: a teórica, cujas conclusões são baseadas na manipulação matemática, enfatizando os dados estatísticos; e a empírica, que descreve a situação estudada sob o ponto de vista do observador, para quem os dados qualitativos são mais importantes.

Essas diferenças fazem com que o modelo teórico da lei não se ajuste aos estudos empíricos, já que esse modelo não permite a possibilidade de inclusão de variáveis que caracterizem a situação real estudada.

Por esse motivo, os estudos empíricos, com aplicação da lei, sem considerar essas variáveis particulares à pesquisa efetuada, correm o risco de chegar a conclusões não confiáveis.

Para que tal risco possa ser evitado ou minimizado, é fundamental que se realizem estudos teóricos mais profundos sobre os processos que resultam na distribuição dos artigos sobre determinado assunto em periódicos, objeto da lei de Bradford. Dessa forma, o modelo teórico poderia se ajustar melhor às situações empíricas.

Por outro lado, conforme se sabe, esse conflito entre teoria e prática é muito comum nas Ciências Sociais, cujo comportamento não é previsível, como nas Ciências Exatas. Isto faz com que os modelos teóricos nos campos ligados àquelas Ciências não permitam refletir efetivamente seus fenômenos, inclusive porque nos dias de hoje ainda não se dispõe de recursos matemáticos necessários para resolver problemas referentes à grande diversificação do comportamento humano. Como a literatura reflete esse comportamento, a matemática torna-se muito limitada ao ser a ela aplicada.

Abstract – Bradford's law is a phenomenon easy to observe in real situations, but its theoretical model does not fit the empirical studies, because the variables of the empirical situation are ignored by the theory. This problem causes errors in conclusions of some studies which can be responsible for bad decisions in terms of management of information systems.

6 Referências bibliográficas

1. PRITCHARD, A. Statistical bibliography or bibliometrics? *Journal of Documentation*, 25: 348, Dec. 1969.
2. BRADFORD, S.C. Sources of information on scientific subjects. *Engineering*. 137: 85-6, Jan. 1934.
3. BRADFORD, S.C. **Documentação**. Rio de Janeiro, Fundo de Cultura, 1961.
4. Idem, p. 13-4.
5. Idem, p. 196-216.
6. VICKERY, B.C. Bradford's law of scattering. *Journal of Documentation*, 4: 198-203, Dec. 1948.
7. GOFFMAN, W. & WARREN, K.S. Dispersion for papers among journals based on a mathematical analysis of two diverse medical literatures. *Nature*, 221(578): 1205-207, Mar. 1969.
8. BROOKES, B.C. Bradford's law and the bibliography of science. *Nature* 224: 953-6, Dec. 1969.
9. PRICE, D.S. A general theory of bibliometric and other cumulative advantage process. *JASIS*, 27: 292-306, Sept./Oct. 1976.
10. DROTT, M.C. Bradford's law: theory, empirism and the gaps between. *Library trends*. Summer 1981.
11. PINHEIRO, L.V.R. **Lei de Bradford: uma reformulação conceitual**. Rio de Janeiro, IBICT/UFRJ, 1982. Dissertação de Mestrado em Ciência da Informação.