

Métodos quantitativos de apoio à bibliometria: a pesquisa operacional pode ser uma alternativa?

Paulo César Rodrigues Borges

Engenheiro cartógrafo e mestre em Sistemas e Computação pelo Instituto Militar de Engenharia (RJ); professor universitário; doutorando do curso de Ciência da Informação do Departamento de Ciência da Informação e Documentação da UnB.
http://pcrb.tripod.com.br

Resumo

O objetivo deste trabalho é apresentar uma forma alternativa para aplicar os métodos da Pesquisa Operacional aos fenômenos bibliométricos que surgiram no início do século XX, até hoje muito polêmicos. Dentre as várias formulações no campo da bibliometria, a chamada "lei de Bradford" foi o foco da investigação. Tentativas deste gênero podem ser uma saída para sistematizar conceitos na bibliometria, confirmando ou descartando descrições e princípios oriundos de suas formulações empíricas. Tendo por base uma linha de analogia entre fenômenos físicos da Teoria do Caos – resolvidos pela Pesquisa Operacional (PO) – e casos de oferta e procura de periódicos, é possível encontrar uma explicação para o comportamento anômalo da curva de Bradford em certas condições críticas. Para aduzir alguma evidência empírica para este ensaio, dois casos práticos na área da PO foram adaptados para a resolução de problemas bibliométricos típicos. Além disso, ao longo de todo o texto, foram assinalados alguns pontos que parecem comuns entre a bibliometria e a Teoria do Caos. Este ensaio, portanto, enseja uma nova questão: a PO poderá contribuir com a ciência da informação, suprimindo-a com modelos determinísticos e bayesianos para explicar os fenômenos bibliométricos?

Palavras-chaves

bibliometria; Lei de Bradford; Pesquisa operacional; Caos; Ciência da informação; Inferência bayesiana.

Quantitative methods to support Bibliometry: can operational research be an alternative?

Abstract

This work aims to present an alternative way to apply the methods of Operational Research to bibliometric phenomena that appeared in the beginning of the XXth century, and that are polemical until now-a-days. Among several formulations in the bibliometric area, the so called "Bradford's law" is the focus of the present research. Such attempts could be a way to fit bibliometric concepts into a reasonable form, sustaining or discarding descriptions or principles derived from its empirical formulæ. Comparing the journal scattering with some similar phenomena that were studied in Chaos Theory – solved by Operational Research (OP) –, it is possible to find an explanation for the anomalous behavior of Bradford's curves in certain critical conditions. In order to bring some evidence forward to this essay on Bibliometry, two practical cases of OP were adapted in order to solve typical bibliometric problems. Furthermore, there are hints throughout this paper that points out commonalities between Bibliometry and the Chaos Theory. Hence, this essay brought to bear a question on us to think about: could OP contribute to Information Science with deterministic and bayesian models to explain bibliometric phenomena?

Keywords

Bibliometry; Bradford's law; Operational research; Chaos; Information science; Bayesian inference.

INTRODUÇÃO

Este artigo surgiu da extensão de um trabalho de pesquisa na disciplina de Pesquisa em Ciência da Informação, no curso de doutorado em Ciência da Informação da UnB (2001), cujo objetivo era analisar os diversos métodos de tratamento dos fenômenos bibliométricos que surgiram no início do século XX, até hoje muito polêmicos. A extensão aludida alinha-se com a proposta de Thomas Kuhn para a gênese de ciências *pós-modernas*, tentando mostrar que é possível produzir conhecimento em bibliometria por analogia com métodos de outros campos do conhecimento aplicados a estes fenômenos.

De uma revisão de literatura expedita de artigos relativamente recentes que tratam da "lei" de Bradford, relativa à dispersão de artigos em diferentes publicações periódicas, surgiu a idéia de investigar os fenômenos bibliométricos por uma linha de analogia aos que foram o embrião da Ciência do Caos, de características *pós-modernas*⁴ como a Ciência da Informação.

O risco que acompanha esta analogia é amenizado de duas formas. Em primeiro lugar, vem a certeza de que a "lei" de Bradford e suas congêneres na bibliometria – "lei" de Zipf e "lei" de Lotka –, trazem em si o problema fundamental de relacionar essas formulações empíricas a teorias científicas, atormentando muitos pesquisadores: Rudolf Carnap e Anatol Rapoport (*apud* O'Connor¹), Drott² e Hubert³. Em segundo lugar, vem a oportunidade de contribuir com a integração dos métodos qualitativos e quantitativos, tão decantada pelos estudiosos de metodologia científica.

O germe da concitação a pesquisas futuras neste assunto de natureza multidisciplinar, repondo em questão às controvérsias bibliométricas, é o que fica de mais expectável. Espera-se que o atrevimento em lançar o pano de fundo da Teoria do Caos e o material empírico da Pesquisa Operacional (PO), como tentativa de compreender a fenomenologia bibliométrica, não seja de vã contribuição.

⁴ Ver noções de ciência *pós-moderna* em Cardoso, A.N.P. *Pós-modernidade e informação: conceitos complementares?* Belo Horizonte: Perspectiva-CL, v.1, n.1, p.63-79, jan./jun. 1996.

AS REVOLUÇÕES CIENTÍFICAS

No final da década de 60, a ciência se encaminhava para uma crise de especialização crescente. Dramaticamente, essa tendência para a especialização foi revertida em virtude do surgimento da Ciência do Caos. Os primeiros teóricos do caos eram sensíveis aos padrões, tinham um gosto pelo aleatório, pelo complexo, pelas extremidades recortadas e pelos saltos súbitos. Pensavam em fazer recuar uma tendência na ciência dita natural – o *reducionismo* –, a análise dos sistemas em termos de suas partes constitutivas. Acreditam estar à procura do todo^{4,5}.

Os sistemas mais simples criam, como hoje se acredita, os mais difíceis problemas de previsibilidade. Não obstante, a ordem surge espontaneamente nesses sistemas – o caos e a ordem, juntos. O estudo moderno do caos começou com a assustadora compreensão, na década de 60, de que equações matemáticas muito simples podiam servir de modelo para sistemas tão violentos, sob todos os aspectos, quanto uma queda d'água^{4,6}. Pequenas diferenças de insumo podiam transformar-se rapidamente em esmagadoras diferenças de resultado – um fenômeno que recebeu o nome de “dependência sensível das condições iniciais”.

Onde começa o caos, a ciência clássica pára. Este adágio traduz o teor das idéias de Thomas S. Kuhn*, que provocaram tanta hostilidade quanto admiração ao serem publicadas em 1962^{4,7}. Ele deu uma boa alfinetada na idéia tradicional de que a ciência progride pela acumulação de conhecimento. Na visão de Kuhn, a ciência normal consiste, em grande parte, em operações de limpeza, partindo de onde alguns cientistas pararam, descartando o que outros descobriram, enfim, fazendo avançar as fronteiras da ciência sem replicação de esforços na geração de conhecimento.

Mas há as revoluções, segundo Kuhn. Uma nova ciência nasce de uma outra, que chegou a um ponto morto. Com frequência, uma revolução tem um caráter interdisciplinar – suas descobertas principais vêm, muitas vezes, de pessoas que se aventuraram fora dos limites normais de suas especialidades. Os problemas que preocupavam esses teóricos não eram considerados linhas de investigação legítimas. Isso era muito frustrante para eles. Todo cientista que se voltou cedo para essas ciências em estado de gestação tem uma história de desânimo ou de hostilidade clara para contar. Propostas de suas teses foram rejeitadas e seus artigos não eram publicados. Idéias

superficiais podem ser assimiladas; idéias que exigem uma reorganização da imagem que se faz do mundo provocam hostilidade.

Casos de interesse para a ciência da informação

Edward Lorenz foi o protagonista de uma dessas revoluções. Ele trabalhou em um computador (*valvulado*) da década de 60, em seu escritório do MIT. Embora esta máquina não tivesse a velocidade nem a memória necessárias para uma simulação realista da atmosfera e dos oceanos terrestres, Lorenz conseguiu representar razoavelmente bem o tempo atmosférico.

No inverno de 1961, querendo examinar mais detidamente uma seqüência, Lorenz tomou um atalho. Em lugar de refazer toda a seqüência, começou pelo meio. Para dar à máquina suas condições iniciais, digitou os números diretamente da impressão anterior. Quando retornou, notou que a nova seqüência não se deu como uma repetição exata da anterior. Isto não era esperado. O programa não fora modificado. Não obstante, ao olhar para a nova impressão, Lorenz viu seu tempo divergindo tão rapidamente do padrão da última seqüência, que, em poucos meses (de laboratório), toda a semelhança desaparecera. Lorenz deduziu que foram os resíduos de cálculo que desprezara até então a causa da mudança. É neste aspecto que um cientista difere de um mero seguidor de fórmulas rígidas de metodologia científica. Lorenz, a partir daí, não desprezaria mais o que era imprevisível no fenômeno complexo da atmosfera, mesmo sendo de pequena influência.

Para se prever o comportamento de um sistema, a tradição científica de Newton até Lorenz determinava que deveriam ser desprezadas pequenas influências, tendo-se à mão um conhecimento sobre as condições iniciais deste sistema e uma lei natural que governasse o fenômeno em estudo, mas, quando Lorenz descobriu (por um computador) um padrão em meio ao ambiente caótico das efemérides meteorológicas, desencadeou um interesse inusitado e crescente em pesquisadores de outras áreas. Por que não aplicar esta metodologia aos fenômenos bibliométricos?

O experimento de Lorenz recebeu um nome técnico: *dependência sensível das condições iniciais*, o que só se dá em um sistema de equações não-lineares, significando isso que expressava relações que não eram rigorosamente proporcionais. As relações lineares são de compreensão fácil e têm solução, o que as torna adequadas para os manuais. Os sistemas não-lineares não podem, em geral, ser solucionados^{4,13}.

* Em seu famoso ensaio: *A Estrutura das Revoluções Científicas*.

A tradição de examinar os sistemas de forma local – de isolar os mecanismos para depois somá-los – estava começando a desmoronar, tanto nas ciências exatas como nas sociais. Kuhn estava certo! Novas ciências estavam a caminho, e o caos parecia inaugurar esta fase. Mas Kuhn era mais um filósofo do que um físico ou matemático. Era preciso que cientistas do embrionário estágio dessa ciência, por experimentos controlados e sistematizados nos *corpora* da física e da matemática, pudessem dar sustentação ao que se vislumbrava em termos de Teoria do Caos. Esses cientistas inauguraram uma forma de conceber toda a complexidade dos sistemas dinâmicos e mesmo de pensar cientificamente*. Físicos, astrônomos e biólogos se convenceram de que tinham de se inteirar das novidades comuns.

Um fato de bastante interesse para a ciência da informação, que mais pode trazer contribuição para este trabalho introdutório, é o relacionado às populações de espécies animais como sistemas dinâmicos, especialmente o cenário maltusiano de crescimento ilimitado de peixes. Os estudos e conclusões do biólogo *Robert M. May* constituem um marco para a Teoria do Caos e um ponto de intrigante questionamento para a formulação dos problemas bibliométricos.

Em resumo, May coletou dados suficientes de um viveiro de peixes e definiu uma equação que podia prever o desenvolvimento (queda e ascensão) dessa população. Para tornar mais realista o experimento, May introduziu um parâmetro que quebrava a linearidade do sistema, sendo possível lidar com as seguintes questões: O que acontece se colocarmos mil peixes em um tanque com um abastecimento limitado de alimentos? O que acontece se juntarmos a isso 50 tubarões que comem dois peixes por dia? O que acontece se um vírus for introduzido nesse ambiente, matando um certo número de peixes a uma certa taxa no tempo, dependendo da densidade populacional?⁴ Os biólogos idealizaram essas perguntas, a fim de poder aplicar as fórmulas rigorosas de modelos matemáticos deduzidos para o caso. E por que não expandir esse tipo de perguntas para o campo da bibliometria? Há mais de sete décadas são feitas as seguintes perguntas: “O que acontece se pusermos uma coleção de 1000 livros em uma biblioteca sujeita a demandas diferenciadas de diferentes tipos de utentes? O que acontece com uma coleção em relação ao seu grau de obsolescência, dadas certas

injunções: evolução dos meios de publicação eletrônica dessa coleção e contingenciamento orçamentário da unidade de informação, por exemplo? Na biologia (e na bibliometria, também), uma simplificação útil foi fazer um modelo do mundo em termos de intervalos de tempo separados. As equações diferenciais descrevem processos que se modificam suavemente com o tempo, mas são difíceis de ser calculadas. Era preciso encontrar uma função linear crescente que constituísse o esquema maltusiano clássico do crescimento da população, não limitado pelo abastecimento de alimentos ou pela contenção moral.

Sem entrar nos pormenores da experiência de May, importa saber que ele levantou uma equação para o seu universo de reprodução de peixes, cujos resultados tabulados mostraram-se tão intrigantes quanto os formulários de efemérides meteorológicas de Lorenz. O algoritmo, facilmente programável em uma pequena calculadora atual, utiliza a população do passo anterior como semente para a próxima interação, verificando-se que a população cresce bastante até uma determinada interação do início do ciclo de cálculo e depois começa a diminuir. Mais adiante no ciclo, os valores começam a oscilar em torno de um valor de convergência e, de repente, a função se estabiliza.

É válido lembrar que, nos dias do lápis e papel e das calculadoras a manivelas, nunca se pôde ir além de poucas interações^{4,5,6}. Na era da microeletrônica e da informática, iniciada virtualmente na década de 70, aqueles cálculos que ficavam na casa de dezenas de iterações puderam ser estendidos a até milhões de iterações, revelando o caos e padrões nele imersos, sendo este aspecto de muita importância para a bibliometria, em particular, e para a ciência da informação, de modo geral, pelo que sugere do todo, isto é, de como surge uma ciência *pós-moderna*.

No experimento de May, a introdução do parâmetro de não-linearidade no seu modelo matemático poderia significar fecundidade dos peixes e outros fatores que poderiam afetar a sua reprodução. A pergunta que esses “novos biólogos” (ecologistas) faziam era a seguinte: “Como esses parâmetros diferentes afetavam o destino final de uma população mutável? A resposta óbvia era que um parâmetro inferior faria com que essa população idealizada terminasse em um nível menor e um parâmetro maior levaria a uma estabilidade maior. Isso se revelou correto para muitos parâmetros – mas nem para todos. Quando parâmetros mais altos foram tentados, o caos começou a surgir. Estranhamente, o fluxo de números começou a se comportar de maneira anômala, o que é um

* Peter M. Senge, autor de *A 5ª Disciplina* (1990), propõe o termo *metanóia* (do gr: *meta* – além – e *nous* – mente): mudança de mentalidade.

aborrecimento para quem está usando uma calculadora manual. Não há convergência nos resultados. De qualquer modo, se a população continuava a ir e vir, os ecologistas supunham que ela oscilava em torno de algum equilíbrio subjacente. O equilíbrio era o importante. Não lhes ocorreu que poderia não haver equilíbrio. Mas para eles só interessavam soluções estáveis. Ninguém queria perder tempo em uma linha de investigação que se tornava irregular, não produzindo nenhuma estabilidade^{4,5,6}.

Então, May passou a testar o seu modelo matemático em condições críticas (no limite), percebendo que os gráficos que retratavam um comportamento previsível sobre a ascensão ou queda de uma população de peixes, depois de muitas interações, apresentavam um comportamento caótico. Não obstante, integrando esses gráficos, May notou regiões (franjas) de regularidade em meio à região caótica, que se repetiam em uma recorrência infinita.

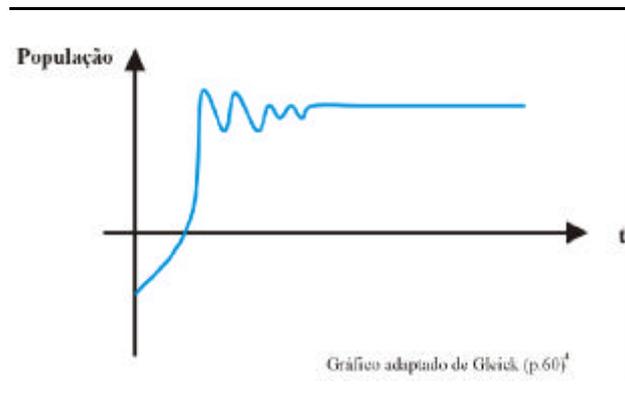
O que fica dessas experiências? As simulações decompõem a realidade em pedaços, tantos quanto possível, mas sempre menos do que o necessário. Um modelo de computador é apenas um conjunto de regras arbitrárias, escolhidas pelos programadores. E o potencial da surpresa do mundo real? É possível formalizar? Sempre que um bom físico examina uma simulação, deve perguntar que parte da realidade foi deixada de lado, que surpresa em potencial foi contornada. A simulação em computador pode desenvolver a intuição, evitar gastos em exercícios ou testes reais de campo, aperfeiçoar os cálculos, mas não dá origem a descobertas autênticas. A intuição do cientista é a chave para as ciências pós-modernas.

PASSANDO UM ESPANADOR NA BIBLIOMETRIA

Lei de Bradford: o “rato de laboratório” para a ciência da informação?

Até aqui, algumas notas de referência neste trabalho tentaram mostrar a similaridade dos eventos bibliométricos e os de outras ciências tributárias da Teoria do Caos. Daqui em diante, os tópicos mais significativos da “lei” de Bradford serão “inspecionados” (e não analisados) nos seus aspectos gerais e para que o leitor capte a intenção de abrangê-la pelo enfoque do caos e pelo método da PO. Não vai aqui nenhuma opção de base formal por esta parte da bibliometria. Não há sustentação para isso, pois nada em termos de cálculo numérico computacional e mesmo desenvolvimentos algébrico e estatístico puros foram efetuados no trabalho, a fim de sustentar esta escolha, que se deu em termos puramente práticos e apegados à analogia, pois, como se viu, a similaridade de siste-

FIGURA 1
Gráfico do regime estacionário de uma população



mas não-lineares estudados pelo caos como a dispersão de periódicos parece ser, intuitivamente, um caminho certo a trilhar. Espera-se que esta explicação seja suficiente para afastar a idéia de uma escolha arbitrária (uma predileção subjetiva) pela “lei” de Bradford*, se forem considerados como *corpus* de estudo bibliométrico as congêneres formulações de Lotka e de Zipf. Suplementando esta explicação, além da similaridade fenomenológica citada, há o fato de que alguns problemas de Pesquisa Operacional também podem ser extensíveis à bibliometria, com pequenas adaptações, como se verá a seguir. Abaixo, os aspectos mais significativos da “lei” de Bradford:

- **Criador da “lei”:** Samuel C. Bradford.
- **Ano da criação da “lei”:** 1934.
- **Problema que suscitou a “lei”:** dispersão dos periódicos nas áreas de geofísica aplicada e de lubrificação.
- **Modelo matemático:**

EQUAÇÃO 1

$$F(x) = a + b \log x$$

Em que $F(x)$ é o número cumulativo de referências contidas no periódico x mais produtivo e a e b são coeficientes. Há uma fórmula parametrizada da “lei” de Bradford, também muito difundida:

* A formulação de Bradford, neste trabalho, não foi considerada como uma lei científica na acepção mais rigorosa do termo, sendo limitada por aspas e iniciada por “l” minúsculo, dando-lhe o cunho de formulação.

EQUAÇÃO 2

$$a = 10^s \quad b = 10^{s+r} \quad g = 10^{s+2r}$$

Em que r é uma unidade de medida fixa no eixo das abscissas e s no eixo das ordenadas. Fazendo $10^s = n$, dividindo-se cada equação do sistema de equações da fórmula anterior por 10^s e renomeando-se os termos, a fórmula transforma-se em uma interessante relação de três membros: **1: n: n²**. Esta relação é muito aplicada no lugar da Equação 2, por admitir a noção de **núcleo de periódicos** – conjunto de periódicos mais especificamente dedicados a um certo assunto. A dispersão ficaria mais categorizada, isto é, haveria grupos ou zonas de periódicos especializados contendo o mesmo número de artigos que o núcleo. A quantidade de periódicos dessas zonas obedeceria à relação **1: n: n²**.

– **Representação gráfica:** ver figura 2.

– **Teor da “lei”:**

Em termos simples, a “lei” diz que não adianta aumentar excessivamente a quantidade de periódicos, porque a soma dos artigos publicados nos periódicos mais importantes (mais consultados) não vai passar de uma quantidade que se estabilizará ou que tenderá a crescer muito pouco, segundo o comportamento de uma função semilogarítmica.

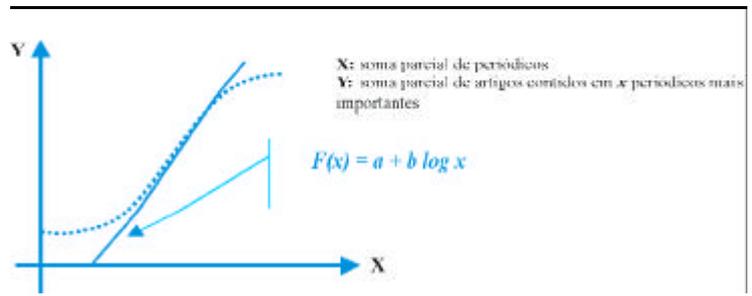
Críticas à “lei” de Bradford

Drott²: desse autor têm-se argumentações mais ligadas aos aspectos teóricos *versus* os empíricos. Ele começou explicando que a formulação de Bradford foi enunciada com base na regularidade observada na recuperação de informações publicadas sob os aspectos de sua concentração e dispersão. Sua crítica se inicia pelas respostas que deixou de apresentar no próprio campo em que se formou – o empírico –, por exemplo: Como determinar o tamanho do núcleo de periódicos? Qual é o valor ótimo para se atribuir ao coeficiente a da Eq. 1, em um certo conjunto de dados? No aspecto teórico, questões desta espécie são apresentadas: Qual é a natureza do processo que causa a distribuição? Quais as relações entre as variáveis* levantadas que podem explicar o porquê da distri-

* Tamanho da amostra, área de especialização do periódico, políticas editoriais,...

FIGURA 2

Representação gráfica da formulação de Bradford.



buição regular de artigos por periódicos? O autor alegou que a vulnerabilidade da formulação de Bradford está na falta de sustentação teórica que o empirismo da fórmula obtida imprime às variáveis que caracterizam a distribuição, uma vez que nenhuma delas é capaz de contribuir na concepção de um modelo teórico rigoroso para descrever os fenômenos bibliométricos.

Carnap e Rapoport apud O'Connor¹: também argumentaram que a distribuição de vários eventos por faixas constantes de frequência só permite prever a ocorrência desses eventos, mas não explica as suas causas. Assinalaram que, apesar de muitas das curvas que descrevem estes tipos de distribuição se aproximarem muito de hipérbolas, nada de teórico fica provado para se afirmar que estas curvas realmente pertencem à classe das hipérbolas.

Consenso crítico: os quatro autores citados concordaram, no entanto, que a “lei” de Bradford constitui um apropriado manancial de utilidade para centros de informação que necessitem de referenciais práticos para operações de aquisição, armazenamento e distribuição de documentos.

Hsieh⁸: investigou artigos que tratavam de Inteligência Artificial (IA) e Sistemas Especialistas (SE) no período de 1976-1987, utilizando dois meios tecnológicos da época: o sistema LISA (Library and Information Science Abstracts) e o sistema Wilsonline. Com base nos resultados bibliométricos obtidos, evidenciou uma produção literária crescente desses assuntos nesse período e um significativo deslocamento de contribuições da IA para a ciência da informação.

Bonitz⁹: contrariou parcialmente Drott², Carnap e Rapoport, que, apesar de admitirem uma fragilidade conceitual na “lei” de Bradford, consideravam-na como um método eficiente de tratamento de certas rotinas de uma unidade de informação, especialmente uma biblioteca tradicional. Bonitz criticou a formulação de Bradford

até mesmo para os poucos casos práticos admitidos por seus colegas. Declarando-se como um estudioso profundo de modelos de distribuição para classificação de eventos probabilísticos, Bonitz refutou a já pequena margem de aplicabilidade da lei de Bradford, apresentando dois efeitos ligados às distribuições assemelhadas à de Bradford, expressos por duas variáveis: o efeito da reordenação (ou da reclassificação) e o efeito do agrupamento. No primeiro efeito, o autor ressaltou o fato de ser muito importante estabelecer parâmetros significativos para modelar o mais fidedignamente possível a distribuição, sendo necessário muito controle sobre eles. Afiançou que Bradford trabalhou sobre parâmetros muito específicos, aos quais denominou “parâmetros individuais”. Bonitz disse que não se pode dizer previamente qual parâmetro é o mais indicado para modelar um fenômeno desse tipo. Isto derruba o poder de predição de fenômenos bibliométricos da lei de Bradford, tirando-lhe o *status* de tributária de uma teoria científica. O outro efeito determina uma variação de curvas em torno da reta ideal de Bradford. O efeito do agrupamento traz em si dois indicadores: a estrutura do acervo documental (jornal, por exemplo) e o perfil do usuário inserido em uma organização em demanda por documentos.

Khavam¹⁰: como Hsieh⁸, trabalhou no afã de determinar a taxa de contribuição da IA para a bibliometria e, dos estudos quantitativos que efetuou sobre artigos publicados em 1988, concluiu que esta contribuição foi muito mais significativa para as ciências naturais do que para as sociais.

Zainab et alii¹¹: na mesma linha de Hsieh⁸ e Khavam¹⁰, concluíram, de um estudo de razoável alcance cronológico (1989-1992) e de extensivo material pesquisado, que existe uma tendência crescente de especialização do pessoal de biblioteconomia nas áreas de SEs e SRIs.

Críticas às críticas da “lei” de Bradford

Depois de uma apreciação geral sobre a recente história da formação da Teoria do Caos, percebe-se uma promissora janela que se abriu para um sem-número de investigações de ordem teórica no campo da bibliometria. As dificuldades por que passou Bradford ao formular suas relações de fundo estocástico e, essencialmente, as críticas às suas relações, à metodologia e às conclusões que tirou fazem parte de todo um natural processo científico que já deu o que tinha que dar. É nítida uma quase-coincidência entre o teor das críticas dirigidas à “lei” de Bradford e o ceticismo crítico de que foram alvo os expoentes do caos. Se um aspecto classificatório puder ser em-

pregado nesta modesta revisão de literatura, pode-se dizer que um foco foi sobre a bibliografia das ciências puramente exatas e que o outro foi sobre a biblioteconomia. O interessante é que tanto O'Connor¹, da área da biblioteconomia, como Gleick⁴, Stevens⁵ e Devaney⁶, da área das exatas, citaram estudos de fenômenos que suscitam o aparente paradoxo da “desordem ordenada”, campo fértil para a novata Ciência do Caos e todas as que dela se utilizem para se formar.

Neste tópico dedicado à crítica das críticas, optou-se por eleger para apreciação a obra que pareceu substanciar a crítica mais completa à “lei” de Bradford: a de Bonitz, tanto pelo aspecto epistemológico que encerrava como pelo analítico-formal. O próprio título da obra – *A false taboo: Bradford* – é de impacto crítico; um tanto acerbo em relação ao conteúdo do artigo, que trata de maneira muito racional e sóbria os pontos fracos da “lei” de Bradford e de seus efeitos. Genericamente falando, Bonitz acreditava na validade da “lei” de Bradford somente em situações simples, quando parâmetros específicos são tomados como base para classificação; e, mesmo assim, a maioria dessas situações não está relacionada com as atividades práticas. Nesse aspecto, ele criticou até os que defendiam a aplicação da “lei” de Bradford para gerir o dia-a-dia de uma unidade de informação como uma biblioteca.

Entrando em domínios epistemológicos, Bonitz disse que freqüentemente alguns teóricos não põem em questão a forma pela qual uma distribuição é formada. Associam um certo fenômeno de seqüência de eventos a formulações matemáticas, esquecendo-se de que, para um matemático, não há valor de finalidade para suas deduções e descobertas. Para um matemático, uma curva pode ser aplicada a um fenômeno físico, como o do declínio radioativo de certo mineral, ou mesmo à distribuição de artigos por periódicos, embora haja diferenças gritantes entre as entidades e processos envolvidos. Bonitz também se utilizou de ilustrações que eventos quotidianos sugerem, para tentar explicar os fundamentos de sua crítica, contradizendo-se no aspecto formal que tanto exalta em suas críticas às formulas de Bradford.

A primeira consideração de ordem crítica que se faz às críticas de Bonitz é que, apesar de já haver um vasto repertório de trabalhos sobre o caos, ele não foi capaz de ligar o fenômeno bibliográfico a outros congêneres de outras ciências, ficando muito restrita a sua conclusão, portanto. Ao tentar trabalhar com agrupamentos de artigos de periódicos para o uso de uma organização, Bonitz notou que a dispersão era muito alta e que o valor da

variável a que chamou de efeito do agrupamento do jornal diminuía. O efeito, segundo ele, refletiria tanto a estrutura do periódico como o perfil dos usuários da organização em apreço. Bonitz queria saber qual era o conjunto de periódicos que melhor atenderia às necessidades informativas de uma organização. Como a dispersão de artigos informativos era muito mais alta em um conjunto de periódicos, o autor deduziu que a linearidade falha pela “lei” de Bradford (*crítica a Bonitz*: é óbvio!), admitindo duas alternativas: 1ª) ou se tem na “lei” de Bradford o caminho único para resolver o problema bibliométrico da dispersão de artigos por periódicos; 2ª) ou se deve ficar na “companhia segura” da realidade, identificando-se parâmetros que melhor se adaptem aos reais processos de dispersão documentária.

Bonitz não poderia ter sido tão infeliz em todo o artigo como o foi nestas duas conclusões maniqueístas. A “lei” de Bradford, pela finalidade em que se tem insistido há décadas, aplicada à vida vegetativa de uma unidade de informação, é de uma trivialidade que compromete até o senso comum como fonte de conhecimento. No entanto, colocá-la como alternativa prejudicada diante da outra alternativa claudicante em termos de possibilidade científica, depois da ascensão do caos, é o mesmo que condenar a bibliometria ao inferno eterno, sem nem lhe dar a oportunidade do purgatório. Bonitz não pensou em testar exaustivamente no tempo as interações da simples formulação de Bradford. Em vez disso, optou por recomendar a escolha de parâmetros ajustados à realidade. Que parâmetros? Por mais que se escolhessem tais parâmetros, o próprio autor argumentou que um salto brusco na curva de Bradford era um sinal de anomalia nessa escolha, que deveria receber atenção. É lícito comparar este caso com uma companhia de soldados que entre em forma pela altura ou pelo peso de seus soldados, não só como analogia ilustrativa, mas também para inferir que, se não apenas uma companhia mas um exército de campanha (dezenas de milhares de soldados) entrasse em forma milhares de vezes por altura e milhares de vezes pelo peso, provavelmente eclodiria um atrator estranho (a seguir) como evidência mínima de ordem em um sem-número de aparentes formações distintas que se dariam após o comando: “em forma!”. Pareceu, também, que Bonitz preocupou-se mais com os processos que com as entidades (objetos e coisas do mundo). Isto é contra-producente pela visão sistêmica. Na abstração de um modelo conceitual de um sistema, dá-se, no início, mais atenção à definição dos objetos, para, depois, estabelecerem-se as relações (lógicas e ontológicas) que sobre eles podem se aplicar.

A bibliometria pode contribuir para “debutar” a ciência da informação no concerto das ciências pós-modernas?

Para os cientistas do caos, a modelagem gráfica (atratores estranhos) de regiões de estabilidade em um sistema turbulento era um motor de informação. E com relação à “lei” de Bradford? Não seria possível transformar o modelo matemático original e suas várias extensões em algoritmos computacionais, de modo a se construir uma teoria mais sólida para a bibliometria? É provável que sim, e um atrator estranho poderá surgir, se as experiências não se ativerem somente a cálculos diretos de fórmulas obtidas empiricamente em um espaço amostral reduzido, como foi o de Bradford, em 1934, tentando mostrar como 395 artigos sobre lubrificantes se dispersavam por 164 diferentes periódicos. É provável, também, que os modelos matemáticos para descrever as relações empíricas ligadas à dispersão de artigos em diferentes publicações periódicas não sejam realmente consistentes e, por isso, lancem dúvidas até sobre a própria essência da bibliometria, em que às vezes é posto em xeque o próprio enquadramento dos fenômenos bibliométricos na realidade empírica, observável e mensurável. Mas não se pode descartá-los. Como se viu para a Teoria do Caos, idéias tiradas do esquecimento fizeram avançar as fronteiras de uma nova ciência.

É preciso inaugurar uma *metanóia* bibliométrica e, talvez, os ingredientes para isso sejam os seguintes: 1) mentalidade criativa e interdisciplinar do grupo de trabalho (biblioteconomia + matemática + ciência da computação); 2) algoritmos simples e elegantes, que implementem as diferentes relações construídas para a dispersão de artigos em periódicos, de forma exaustiva; 3) parque tecnológico adequado (*hardware* e *software*); e o item mais caro de obter-se: 4) dados.

PESQUISA OPERACIONAL: UM MÉTODO QUANTITATIVO E ALTERNATIVO DE APOIO À BIBLIOMETRIA

Enquanto não surgir um consenso na definição de uma teoria bibliométrica para as formulações de Bradford e de outros pesquisadores, é conveniente que a estas formulações sejam incorporados os modelos matemáticos de uma disciplina que se firmou no meio científico atual: a Pesquisa Operacional.

A metodologia aplicada da PO seria uma solução intermediária e nova para aumentar o poder de resolução dos problemas bibliométricos, de manter a bibliometria em foco dentro das iniciativas de pesquisa em ciência da

informação e biblioteconomia, podendo suscitar muitos trabalhos finais nos cursos de graduação nessas áreas, e, finalmente, a aplicação dessas técnicas, com todo o aporte computacional que trazem em si, poderia também renovar o interesse do público profissional que trabalha nas unidades de informação modernas.

Generalidades sobre a Pesquisa Operacional

Definir a PO como disciplina científica tem sido uma tarefa extremamente difícil, não só por ser uma disciplina nova, mas também pelo seu caráter multidisciplinar. Grande parte das definições simplistas relegam o papel da PO a um ramo da matemática aplicada, preocupada apenas com problemas de “otimização”. Como simples usuária de ferramentas matemáticas aplicadas a problemas da vida real, reduz-se o seu espectro de aplicações e prejudica-se a sua real posição de disciplina independente. É também um erro muito comum confundir a com a função gerencial.

Há várias definições, mas uma definição integradora, encontrada na disciplina de Teoria da Decisão do Curso de Altos Estudos Militares da Escola de Comando e Estado-Maior do Exército Brasileiro, é a seguinte:

“Pesquisa Operacional é uma pesquisa aplicada e dirigida, dentro das operações de um sistema ou de uma organização, considerados em seu aspecto total. É, outrossim, um método científico de trabalho, que utiliza técnicas e instrumentos científicos, tendo em vista obter o melhor rendimento possível no funcionamento de uma organização ou o melhor desempenho possível de um sistema em estudo. Baseia-se em análises quantitativas para auxiliar o processo de decisão.”

O berço da PO, no Brasil, remonta à década de 50, em que professores das áreas de matemática, estatística e engenharia do IME (Instituto Militar de Engenharia - RJ), ITA (Instituto Tecnológico da Aeronáutica - SP), USP e da UFRJ já se entusiasmavam com o potencial desse campo de pesquisa. Apesar das dificuldades típicas da introdução de novas técnicas e de mudanças de padrões funcionais em uma organização, os resultados da PO têm motivado muitos indivíduos e grupos, acostumados com a rápida absorção de novas tecnologias que dão certo, o que, inevitavelmente, criará um ambiente favorável para a sua disseminação nos meios governamental, comercial e industrial, visto que, no acadêmico, isto já se deu.

O objetivo geral da PO é o de descobrir regularidade em algum fenômeno e ligar essa regularidade com outros conhecimentos, de tal forma que o fenômeno possa ser modificado ou controlado. Não são do escopo da PO atividades de limitar técnicas e instrumentos para a solução de problemas como um todo. Se um problema pode ser resolvido somente por princípios e técnicas de engenharia, então é um problema de engenharia, e não de PO.

A PO, a cada dia que passa, tem sido mais aplicada na área das ciências sociais, especialmente na administração, promovendo um desenlace com o tradicional método quantitativo, monopólio das ciências exatas pela visão *reducionista* do positivismo. O enfoque qualitativo foi trazido, por conseguinte, pelos administradores de empresa, incomodados com o excesso de rigor matemático dos métodos de PO e, muitas vezes, frustrados pela pouca flexibilidade dos modelos, que respondiam somente a perguntas padronizadas. O enfoque qualitativo da PO é o reconhecimento de que o tratamento quantitativo dos problemas fornece uma estrutura de raciocínio e análise que permite descobrir qual é a informação necessária para se desencadear a fase qualitativa de resolver o problema mais amplo. É a *metanóia* na PO.

Em que pese a profusão de variáveis para resolver problemas nesses novos campos de aplicação da PO, uma limitação ao seu método continua a ser o de restringir o número de variáveis para se poder construir um modelo que funcione. A evolução para uma aplicação qualitativa na PO não pôde absorver integralmente a abertura a um número excessivo de variáveis para representar o mundo real. Foi uma troca com limites, mas que, se não resolve todos os problemas desses novos campos, pelo menos acende um farol no fim de um túnel de nebulosos e confusos contextos, que havia antes dessa integração da PO às ciências sociais.

As técnicas de PO mais empregadas estão organizadas nas seguintes categorias: 1ª) Programação Linear: envolvendo problemas de programação linear, programação linear inteira, problemas de transporte, problemas de transbordo, problemas de designação e problemas de alocação de recursos; 2ª) Análise de Redes: envolvendo problemas de árvore geradora mínima, caminho mais curto e fluxo máximo; 3ª) Teoria dos Estoques; 4ª) Teoria das Filas; 5ª) Simulação de Sistemas; 6ª) Teoria dos Jogos; 7ª) Programação Dinâmica: trata de problemas ligados à Teoria da Decisão e da Teoria da Utilidade. Pelo tom que seu deu a este trabalho, só serão utilizadas a 1ª e a 7ª técnicas. É por isso que elas serão descritas a seguir.

A **programação linear** cuida dos problemas de determinação de valores máximos e mínimos de funções, sem contudo tratar de variáveis que produzam ou armazenem valores negativos. Normalmente, o sistema de equações é indeterminado, isto é, há mais variáveis que equações, acarretando um sem-número de soluções para este sistema, que expressa as restrições tecnológicas ou econômicas na liberdade de ação de uma organização. Ainda dentro da programação linear, uma classe de problemas de amplo espectro de aplicação é a de transporte. Um problema de transporte se resume em determinar o carregamento de uma rede de transporte que liga várias fontes a vários destinos, de forma que o custo total do transporte na rede seja mínimo.

Os problemas de **programação dinâmica** estão calcados na disciplina de Probabilidade e Estatística, bem como em um rebento desta: a Inferência Bayesiana da IA. O problema mais essencial desses campos é o de tentar “medir” a incerteza que afeta um sistema, em seus mais variados aspectos. Nesse ponto é bom distinguir algumas noções que diferenciam a análise bayesiana dos métodos probabilísticos convencionais.^{12, 13}

Os problemas de decisão (Teoria da Decisão) exploram quatro tipos de contexto, na ordem crescente de indeterminação: a certeza, o risco, a incerteza e o conflito. As técnicas (V. problema nº 2, a seguir) que podem ser aplicadas a esses contextos são as seguintes: para problemas de certeza, a programação linear; para os de risco, a Teoria da Probabilidade; para os de incerteza, a análise bayesiana; e para os de conflito, a Teoria dos Jogos. A árvore de decisão constitui a ferramenta básica de solução desses problemas. É uma estrutura ramificada ou um diagrama de probabilidades, em que a cada ramo da árvore é associada uma probabilidade de ocorrência de um evento; a cada nó de probabilidade (simbolizado por um círculo) é associado o valor da esperança matemática acumulada para aquele nó; a cada nó de decisão (simbolizado por um quadrado), é associado um valor referente ao somatório mais favorável das esperanças matemáticas dos ramos tributários. O caminho da árvore (de trás para a frente) que maximiza a vantagem ou minimiza a desvantagem constitui a linha de ação a ser adotada. Pode ser necessário aplicar um método estatístico convencional ou mesmo o Teorema de Bayes em algum ramo da árvore, nesses tipos de problema.

A chamada **Inferência Bayesiana**, muito explorada hoje em dia pela IA, vem de muito tempo atrás – do abade e matemático inglês Thomas Bayes (1702–1761). Bayes tentou determinar a probabilidade das causas por inter-

médio dos efeitos observados. *Grosso modo*, o tamanho e conhecimento que se tem da amostra é o fator que determinará a escolha de uma distribuição da Probabilidade e Estatística clássica (funções de verossimilhança) ou do Teorema de Bayes (distribuição *a priori*). Infelizmente, muitas vezes é difícil coletar e tratar dados para uma experiência, nada melhor que utilizar as funções de verossimilhança e trabalhar com estimadores clássicos da Probabilidade e Estatística. Do contrário, se os únicos dados disponíveis são de uma amostra pequena e pouco confiável (mais subjetiva), então é útil trabalhar com uma distribuição *a priori*. Em alguns problemas de ciências sociais, a estimativa com base em grandes amostras em que se pode confiar é impossível ou proibitivamente dispendiosa. Então, a análise bayesiana proporciona um esquema para a tomada de decisões racionais¹³. O “velho” Teorema de Bayes, na atual conjuntura de *pós-modernidade*, acabou por produzir fecundos rebentos em pesquisas na área de IA, voltadas à avaliação do conhecimento impreciso (incerteza). Daí veio a Teoria da Evidência de A. P. Dempster (1968), posteriormente reformulada por G. Shafer (1973), cujos objetivos confluem para a formalização de uma “metodologia de decisão”, baseada em uma função de crença, em uma função de perda e no princípio da maximização da utilidade esperada, em um contexto de ignorância parcial, sendo vasto o campo de aplicações, por exemplo: 1) preferências eleitorais; 2) indexação automática de documentos; 3) representação da incerteza de regras, solução de conflitos de regras, etc.¹²

Resolução de problemas bibliométricos por programação linear e por árvore de decisão

Para a elaboração de um modelo de solução para um problema de PO, os passos são os seguintes: 1º) formulação do problema; 2º) definição das variáveis de decisão; 3º) estabelecimento da função-objetivo; 4º) formulação das equações de restrição; 5º) formulação do modelo. Neste tópico, as fases acima serão vistas em situações reais da bibliometria, adaptadas de problemas de PO na área de administração de empresas.

Problema nº 1 (programação linear):

O bibliotecário-chefe de uma grande unidade de informação está tentando “otimizar” as diversas aquisições de periódicos que faz a quatro tipos de usuários: universitários, industriais/comerciantes, estudantes de 2º grau e estudantes de 1º grau. Os recursos de que dispõe para isso permitiriam a aquisição de 50.000 exemplares de quatro categorias correspondentes àqueles públicos-al-

vos. As probabilidades de sucesso e de riscos com os empréstimos correspondentes a cada categoria foram tabuladas como mostra a tabela 1. Há um mínimo de restrições que devem ser seguidas:

1ª) a aquisição de periódicos para estudantes do 1º e 2º graus não deve ultrapassar 40% do valor total de exemplares a serem adquiridos;

2ª) as aquisições para estudantes do 1º grau não podem exceder 20% do total dos exemplares a serem adquiridos para estudantes do 1º e 2º graus;

3ª) o total das aquisições para industriais/comerciantes não deve exceder a 10.000 exemplares;

4ª) a probabilidade de risco não deve exceder 0,8% da aquisição total.

Pede-se determinar, por programação linear, o total de exemplares a serem adquiridos em cada tipo de público, de modo a maximizar o sucesso na operação de empréstimo de periódicos dessa grande unidade de informação.

Solução:

a. Definição das variáveis de decisão

- X_1 = aquisição para universitários
- X_2 = aquisição para industriais/comerciantes
- X_3 = aquisição para estudantes do 2º grau
- X_4 = aquisição para estudantes do 1º grau

b. Estabelecimento da função-objetivo

Com se deseja que o sucesso na aquisição seja máximo, deve-se considerar os públicos de maior e menor risco para a aquisição. Assim, se existe 0,2% de risco na aquisição de periódicos para os universitários, há 99,8% de possibilidade de sucesso. Logo, para a primeira categoria de aquisição, tem-se $(0,998 \times 0,09 - 0,002)X_1 = 0,08782X_1$

De forma análoga, para as outras categorias, tem-se:

$$(0,995 \times 0,10 - 0,005)X_2 = 0,0945X_2$$

$$(0,990 \times 0,13 - 0,010)X_3 = 0,1187X_3$$

$$(0,980 \times 0,14 - 0,020)X_4 = 0,1172X_4$$

TABELA 1
Probabilidades de sucesso e risco no empréstimo de periódicos.

Usuário	Probabilidade de sucesso	Probabilidade de risco
Universitário	9%	0,2%
Industrial/comerciante	10%	0,5%
Estudante do 2º grau	13%	1,0%
Estudante do 1º grau	14%	2,0%

Então, a função-objetivo será a seguinte: Maximizar $z = 0,08782X_1 + 0,0945X_2 + 0,1187X_3 + 0,1172X_4$

c. Formulação das inequações de restrição

- As aquisições para os estudantes do 1º e do 2º graus não devem ultrapassar 40% do total: $X_3 + X_4 \leq 0,4 (X_1 + X_2 + X_3 + X_4)$, ou seja: $-0,4X_1 - 0,4X_2 + 0,6X_3 + 0,6X_4 \leq 0$

- As aquisições para os estudantes de 1º grau não devem exceder 20% do total de aquisições para estudantes do 1º e 2º graus: $X_4 \leq 0,2 (X_3 + X_4)$, ou seja: $-0,2X_3 + 0,8X_4 \leq 0$

- O total de aquisições para os estudantes do 1º grau não deve exceder a 10.000: $X_2 \leq 10.000$

- As aquisições para os universitários não devem ser menores que 20.000: $X_1 \geq 20.000$

- O risco não deve ultrapassar 0,8% das aquisições totais:

$$0,002 X_1 + 0,005X_2 + 0,01X_3 + 0,02X_4 \leq 0,008 (X_1 + X_2 + X_3 + X_4)$$

Fazendo as operações e simplificações necessárias, vem: $-0,6X_1 - 0,3X_2 + 0,2X_3 + 1,2X_4 \leq 0$

- A quantidade total de aquisições é de 50.000, donde vem que $X_1 + X_2 + X_3 + X_4 \leq 50.000$

Pelo princípio da não-negatividade da programação linear, tem-se que $\{X_1, X_2, X_3, X_4\} \geq 0$

d. Formulação do modelo

Encontrar valores para $\{X_1, X_2, X_3, X_4\}$, de forma a maximizar a função-objetivo já estabelecida, com as restrições levantadas na letra c. Como o número de equações é maior que o número de incógnitas (variáveis), o

sistema é superabundante, admitindo infinitas soluções. A maneira de se resolver este tipo de problema não será aqui explicitada e foge ao objetivo deste trabalho (não faz parte dos métodos normais de solução da álgebra linear), valendo dizer que foi empregado um programa de ajustamento complexo (transformação estatística), desenvolvido na linguagem de programação Fortran 77. Daí, o resultado obtido em números de exemplares foi o seguinte:

$$X_1 = 20.000; X_2 = 9.999; X_3 = 20.000; X_4 = 0$$

Problema nº 2 (programação dinâmica – árvore de decisão):

Neste problema não se fará uma adaptação específica a um problema bibliográfico. Ele será expresso em uma forma o mais genérica possível, de tal sorte que possa ser adaptado a rotinas administrativas de uma organização para a aquisição de bens ou para o oferecimento de serviços o mais adequados e vantajosos possíveis.

Enunciado: suponha que uma organização vá lançar um produto ou oferecer um serviço para um universo de usuários (cativos, potenciais etc.). A diretoria da organização reúne-se e levanta três linhas de ação:

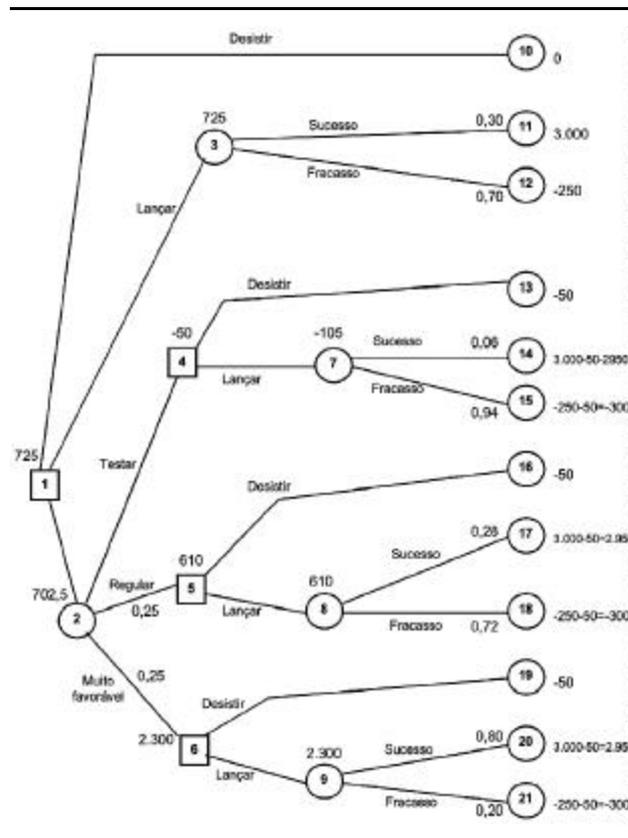
- a) desistir do lançamento do produto ou da prestação do serviço;
- b) lançar imediatamente o produto ou prestar imediatamente o serviço;
- c) efetuar uma pesquisa de mercado (produto) ou distribuir um questionário (serviço) para optar por A ou por B, dependendo do resultado.

Os resultados das pesquisas podem ser “pouco favorável”, “regular” ou “muito favorável”, com as probabilidades de 50%, 25% e 25%, respectivamente. A probabilidade de o produto ou o serviço fazerem sucesso sem a pesquisa é de 30% (note que as probabilidades já foram fornecidas, não sendo necessário calculá-las).

O departamento de planejamento da organização avaliou que o sucesso propiciará um retorno de 3.000 UM (unidades monetárias) e o insucesso, um prejuízo de 250 UM. Sabe-se que o custo da pesquisa será de 50 UM.

Conhecem-se, ainda, as probabilidades de o produto ou de o serviço terem sucesso, dado que a pesquisa apresentou os seguintes resultados:

FIGURA 3
Árvore de decisão do problema nº 2



- “pouco favorável” = 6%
- “regular” = 28%
- “muito favorável” = 80%

Pede-se desenhar a árvore de decisão do problema e determinar a melhor linha de ação para a organização, baseando-se no critério do valor matemático esperado.

A Eq. 3 é a fórmula para o cálculo do valor matemático esperado.

O coeficiente c_i é o valor matemático esperado do evento e $P(c_i)$ é a probabilidade associada à ocorrência.

Solução:

- Da árvore de decisão da figura 3, extraem-se os seguintes valores:

$$VME_3 = 3.000 \times 0,30 - 250 \times 0,70 = 725$$

$$VME_7 = 2.950 \times 0,06 - 300 \times 0,94 = -105$$

$$VME_8 = 2.950 \times 0,28 - 300 \times 0,94 = 610$$

$$VME_9 = 2.950 \times 0,80 - 300 \times 0,20 = 2.300$$

$$VME_2 = -50 \times 0,50 + 610 \times 0,25 + 2.300 \times 0,25 = 702,5$$

– Decisão: como $VME_3 > VME_2$, a melhor alternativa é lançar o produto ou oferecer o serviço, sem realizar a pesquisa.

$$VME_i = \sum_{j=1}^n c_j P(c_j)$$

CONCLUSÃO

É bom lembrar que na recente história das disciplinas que compõem o *corpus* da ciência da informação, a eventual contribuição de certos “forasteiros” das ciências ditas “duras” (*hard sciences*) ou físicas foi muito auspiciosa. O austríaco Eugênio (Eugen) Wüster, que desenvolveu a Teoria Geral da Terminologia, foi um exemplo, pois era um engenheiro. Benoit Mandelbrot, matemático e um dos “pais” da Teoria do Caos, interessou-se muito pela enigmática “lei” de George Kingsley Zipf^{*}, filólogo de Harvard. Portanto, mais uma vez, fica o convite para que mais “forasteiros” se interessem em destrinçar essas formulações empíricas que foram orladas desde o início do século por pesquisadores da biblioteconomia, da lingüística, enfim, das *soft sciences*, ainda aguardando o toque final da parceria interdisciplinar com as *hard sciences*, para erigir uma disciplina ou mesmo um campo científico, mandamento capital da pós-modernidade científica.

Com a recente história da bibliometria, pode-se dizer que existe uma linha de similaridade com o que ocorreu com o caos. Muitas das iniciativas de pesquisadores do início do século, que deduziram expressões bibliométricas empíricas, foram deixadas de lado, mas podem ser o embrião de uma nova disciplina, se ressuscitadas com novos recursos, principalmente os que a ciência da informação vem carreando para o seu *corpus*, enriquecido por métodos e instrumentos de outras ciências.

Uma conclusão é uma asserção. No entanto, com laivos de intrigar o leitor, prefere-se “questionar” neste item em vez de concluir. As questões que se seguem trazem em si o espírito afirmativo das respostas (seu conjunto = conclusão), pelo objetivo não formalmente expresso deste trabalho de aliar o método seguido pelos criadores do caos ao caminho incerto de a bibliometria descrever os fenômenos de seu interesse. Responder a estas perguntas com incursões investigativas formais é justificar um *constructo* matemático para a bibliometria, norteando a sua análise por métodos validados cientificamente. Sendo assim, eis as perguntas:

– Poderia o aprofundamento da pesquisa bibliométrica trazer à tona uma gama de resultados que a alçasse como disciplina científica?

– Isto seria um cartão de entrada da ciência da informação para o universo das ciências *pós-modernas*?, idealizado por Thomas Kuhn?

– O quadro da não-linearidade é extensível aos fenômenos de dispersão de artigos e outros de interesse da bibliometria?

– Como encarar a não-linearidade bibliométrica, se ela ocorrer? Procurar por um ponto em comum com um fenômeno que suscite um comportamento não-linear de um sistema de outro campo científico é um caminho a ser seguido?

Crê-se que a formalidade que for impressa à resposta dessas questões poderá amenizar as críticas dos autores que, com razão, não enxergam rigor científico nas proposições e modelos matemáticos de Bradford. Para Kuhn, alguns cientistas (sociais ou das exatas) se desorientam quando aplicam seus métodos tradicionais de pesquisa. Eis o momento para que os cientistas da informação aproveitem os insumos metodológicos de outras ciências. A matemática aplicada e o cálculo numérico^{*}, como principais tributárias da PO, podem levar as críticas ao fim ou a planos mais objetivos de orientação. A profissão de cientista tem de encarar as anomalias em seus resultados, não tentando mais conduzi-los à “normalidade” por métodos de aproximações, isto é, desprezando resíduos não-lineares. A Teoria do Caos e os métodos da PO para a solução de problemas caóticos são, por enquanto, os únicos meios disponíveis para lidar com essas “anomalias”.

^{*} Disciplina da ciência da computação que se preocupa com métodos numéricos para resolver, principalmente, problemas matemáticos de divergência na solução de sistemas de equações.

^{*} Esta lei trata da frequência de ocorrência de palavras em um texto.

Métodos quantitativos de apoio à bibliometria: a pesquisa operacional pode ser uma alternativa?

“Lei” de Bradford + técnicas de PO + Teoria do Caos! Talvez repouse aí o pacote de surpresas que dará um novo sentido à ciência da informação, começando por resolver os problemas bibliométricos, ou, melhor que resolvê-los, criar novos problemas nessa estacionária disciplina, que guarda tanto potencial teórico-prático.

Artigo aceito para publicação em 03-12-2002

REFERÊNCIAS

1. O'CONNOR, Daniel O.; VOOS, Henry. *Empirical laws, theory construction and bibliometrics*. Nova Jérsei : Escola de Biblioteconomia e de Estudos de Informação da Universidade de Rutgers University, 1981. p. 9-19.
2. DROTT, M. Carls. *Bradford's law: theory, empiricism and the gaps between*. Filadélfia : Escola de Biblioteconomia e de Ciência da Informação da Universidade de Drexel, 1981. p. 41-51.
3. HUBERT, John J. *General bibliometric models*. Ontário : Universidade de Guelph, 1981. p. 65-79.
4. GLEICK, James. *CAOS: a criação de uma nova ciência*. Rio de Janeiro : Campus, 1990. 310 p.
5. STEVENS, R. T. *Fractal: programming in C*. California : M&T Books, 1989. 583 p.
6. DEVANEY, R. L. *Chaos, fractals, and dynamics: computer experiments in mathematics*. California : Addison-Wesley, 1990. 181 p.
7. ATAÍDE, M. E. M. O lado perverso da globalização na sociedade da informação. *Ciência da Informação*, Brasília, v. 26, n. 3, p. 268-270, set. 1997.
8. HSIEH, Cynthia C. Survey of artificial intelligence and expert systems in library and information science literature. *Information Technology and Libraries*, Califórnia, v. 8, n. 2, p. 209-214, Jun. 1989.
9. BONITZ, Manfred. A false taboo: Bradford. *International Forum on Information and Documentation*, Moscou, v. 16, n. 2, p. 15-17, Apr. 1991.
10. KHAVAM, Yves. L'apport de la bibliometrie aux recherches interdisciplinaires. Le cas de l'intelligence artificielle. *Documentaliste Sciences de L'Information*, Paris, v. 29, n. 3, p. 129-135, mai./juin. 1992.
11. ZAINAB, Awang Ngah; SILVA, Sharon Manel de. Expert systems in library and information services: publication trends, authorship patterns and expressiveness of published titles. *Journal of Information Science*, Brighton, v. 24, n. 5, p. 313-336, 1998.
12. SILVA, Wagner Teixeira. *Algoritmos para raciocínio evidencial usando funções de crença*. 1991. 214 f. Tese (Doutorado) - PUC-RJ. Rio de Janeiro, 1991.
13. WONNACOTT, Thomas H.; WONNACOTT, Robert J. *Introdução à estatística*. São Paulo : LTC , 1980. 589 p.