

書目計量學

Lecture 04 -- 布萊德福經驗法則

陳光華
國立台灣大學圖書資訊學系
國立台灣師範大學社會教育學系
khchen@ntu.edu.tw

大綱

- 布萊德福經驗法則的產生背景
- 布萊德福經驗法則的形成
- 布萊德福經驗法則的基本原理
- 布萊德福經驗法則的理論發展
- 布萊德福經驗法則的應用

布萊德福

- *Samuel Clement Bradford, 1878-1948*
 - 著名的文獻學家和化學家
 - 創立布萊德福文獻分散經驗法則，是文獻計量學的主要奠基人
 - 多產作家，著有大量關於分類理論與實踐、編目理論的著作。
 - 1948年完成鉅作-《文獻學》
 - 文獻分散經驗法則的系統論述

布氏經驗法則的產生背景

- 文獻的分散是普遍的客觀現象
 - 一個學科的論文分散在其他學科的刊雜誌上屢見不鮮
 - 如何找出其分散的規律性是關鍵所在
- 科學統一性原則
 - 每一學科都或多或少，或遠或近地與其他任何一個學科相關聯。
 - 對一個專家有用的論文，不僅會出現在這個專家所在學科的專業期刊上，亦會出現在“其他學科”的期刊上。
 - “其他學科”期刊的數量，和其研究領域與“專家所在”學科的關係密切程度以及“專家所在”學科的論文在每種期刊中的登載率呈反比。
 - 按照某個學科論文期刊的載文率之高低來劃分區域，每個區域中的期刊數量隨著該區域期刊載文率的減小而增多。
- 文獻統計研究是布氏經驗法則產生的基礎

布氏經驗法則的提出

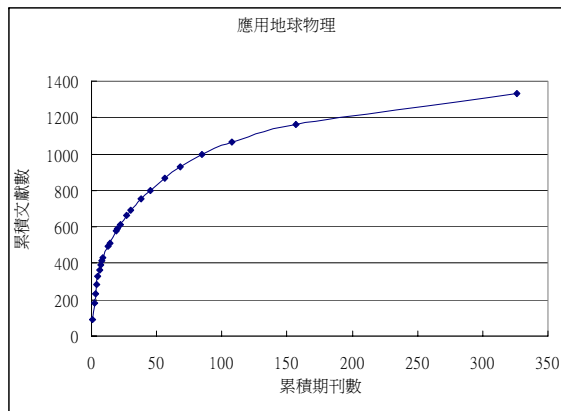
- 樣本：應用地球物理學和潤滑學領域的期刊
- 區域分析
 - 每種期刊每年至少刊載4篇以上相關主題的文章
 - 每種期刊每年平均刊載1~4篇相關主題的文章
 - 每種期刊每年至少刊載1篇相關主題的文章
- 發現
 - 各區相關論文數大體相同
 - 相繼各區的期刊數基本上成等比數列

分區	期刊載文數量	應用地球物理學		潤滑學	
		期刊量	論文章	期刊量	論文章
1	>4	9	429	8	110
2	1~4	59	499	29	120
3	1	258	404	127	152

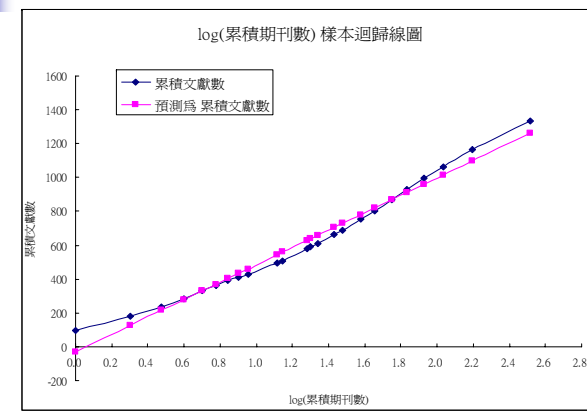
應用地球物理學期刊

期刊數	文獻數	累積期刊數	累積文獻數	log(累積期刊數)	累積文獻數
1	93	1	93	0.0000	93
1	86	2	179	0.3010	179
1	56	3	235	0.4771	235
1	48	4	283	0.6021	283
1	46	5	329	0.6990	329
1	35	6	364	0.7782	364
1	28	7	392	0.8451	392
1	20	8	412	0.9031	412
1	17	9	429	0.9542	429
4	16	13	493	1.1139	493
1	15	14	508	1.1461	508
5	14	19	578	1.2788	578
1	12	20	590	1.3010	590
2	11	22	612	1.3424	612
5	10	27	662	1.4314	662
3	9	30	689	1.4771	689
8	8	38	753	1.5798	753
7	7	45	802	1.6532	802
11	6	56	868	1.7482	868
12	5	68	928	1.8325	928
17	4	85	996	1.9294	996
23	3	108	1065	2.0334	1065
49	2	157	1163	2.1959	1163
169	1	326	1332	2.5132	1332

應用地球物理學期刊 (續)



應用地球物理學期刊 (續)



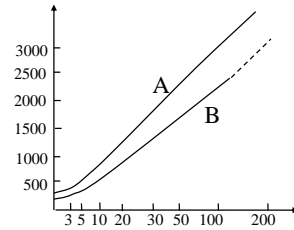
布氏經驗法則的圖像觀察

■ 方式

- 橫軸：期刊累積數
- 縱軸：文章累積數

■ 發現

- 忽略左下部彎曲的部分，為一條直線
- 三個分區的期刊數量亦成等比關係



A: 應用地球物理學 (1929-1932年)
B: 潤滑學 (1931-1932)

布氏經驗法則的推導

m_1, m_2, m_3 : 分別為第1、2、3各區期刊刊載的論文數

- n_1, n_2, n_3 : 相應各區的期刊數
- r_1, r_2, r_3 : 相應各區每種期刊平均刊載的相關論文數
 - $r_1 = m_1/n_1, r_2 = m_2/n_2, r_3 = m_3/n_3$
- 各區域論文數相等， $m_1 = m_2 = m_3$

$$n_1 r_1 = n_2 r_2 = n_3 r_3 = m_1$$

$$\rightarrow n_2/n_1 = r_1/r_2 = a_1, \quad n_3/n_2 = r_2/r_3 = a_2$$

$$\text{又 } a_1 = a_2 = a$$

$$n_2 = a_1 n_1 = a n_1$$

$$n_3 = a_2 n_2 = a_2 a_1 n_1 = a^2 n_1 = a^2 n_1$$

$$\text{則 } n_1 : n_2 : n_3 = 1 : a : a^2$$

布氏經驗法則的確立

■ B. C. Vickery 的對布萊德福經驗法則的研究

- 將布萊德福的研究結論稱作“布萊德福分佈經驗法則”
- 提出自己的修正和補充，使布氏經驗法則更加完整和豐富。
- 布萊德福經驗法則的確立與成功，其厥功甚偉

■ F. F. Leimkuhler

- 對布氏經驗法則之區域描述作了重要發展

■ B. C. Brookes

- 以數學公式描述布氏經驗法則，發展了圖像分析方法。

布氏經驗法則的基本闡述

- 區域描述
- 圖像描述
- 區域描述和圖像描述之比較

區域描述

- 區域分析
 - 將某一時期之接某科學載文量等級排列的期刊劃分為三區，每一區所包含的相關論文數量相等。
- 發現
 - 第一區（核心區）：數量不多但效率最高的 p_1 種期刊
 - 第二區（相關區）：數量較大，效率中等的 p_2 種期刊
 - 第三區（邊緣區）：數量最大而效率最低的 p_3 種期刊
 - 三個區域中的期刊數量關係
 - $n_1 : n_2 : n_3 = 1 : a : a^2$
 - a 為布萊德福乘數（布式係數）

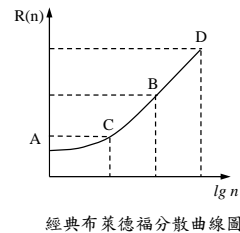
區域描述 (續)

- 布氏經驗法則表示

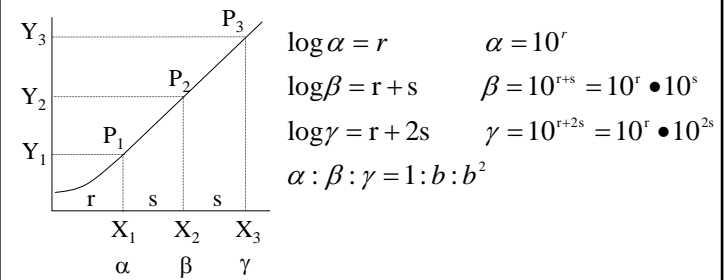
分區	期刊量	論文章
1 (核心區)	8	220
2 (相關區)	40 = 8*5	220
3 (邊緣區)	200 = 8*5 ²	220
- 發現
 - 各區的期刊量，逐區上升
 - 每種期型的載文量逐區下降
 - 核心區資訊密度最高，每種期刊平均刊載27.5篇論文
 - 相關區次之，每種期刊平均刊載5.5篇論文
 - 邊緣區最低，每種期刊平均刊載1篇論文

圖像描述

- 圖像表示
 - 取期刊數量的對數 ($\lg n$) 為橫座標，相應的論文累積數 $R(n)$ 為縱坐標進行圖像描述，得到下圖所示之曲線。
- 圖像描述的另一個結論
 - $n_1 : (n_1 + n_2) : (n_1 + n_2 + n_3) = 1 : b : b^2$
 - $n_{1-2} = n_1 + n_2, n_{1-3} = n_1 + n_2 + n_3$
 - $\rightarrow n_1 : n_{1-2} : n_{1-3} = 1 : b : b^2$
 - n_1 ----- 核心區的期刊數量
 - n_{1-2} ----- 核心區和第二區中的期刊累積數
 - n_{1-3} ----- 全部三個區中的期刊累積數，即期刊總數 (N)
 - b ----- 分散系數 (與前式中的 a 不同)



圖像描述 (續)



區域描述和圖像描述之比較

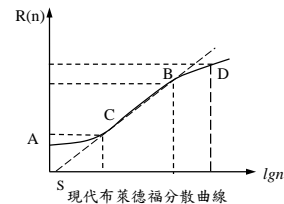
- 區域描述和圖像描述之結論不一致
 - 經過數學證明布氏經驗法則的兩種形式存在著不能統一的矛盾。
- 從數學和實際觀點來看
 - 圖像描述比區域描述更接近於文獻的實際分布情況
 - E. A. Wilkinson之證明
- 就文獻計量學的規律說明
 - 不論是區域描述或圖像描述，皆只能近似地揭示文獻分布的規律，而無法精確地符合統計數據。

布氏經驗法則的理論解釋

- 現代科學技術交叉滲透，小學科又有向大學科綜合的趨勢，產生文獻分布中的“累加效應”。
- “限定因素”之作用
 - 曲線與直線分界點的產生
 - 抑制核心期刊數量之無限增加，此外期刊的篇幅亦有限，編輯人員不得不制定出版計劃和方針，控制文獻的數量
 - 隨時間的流逝，刊登該學科文獻的期刊種數與相關論文數成正比增加，使得布氏經驗法則的圖形呈直線增長。

格魯斯下垂

- 格魯斯 (Q.V. Groos) 認為布萊德福經驗法則之圖形應分為三部分
 - 上升的曲線部分-直線部分-彎曲下垂部分
 - BD之下垂部分稱為格魯斯下垂
- 加拿大波普 (Adrew Pope) 進行統計分析亦證明圖形構成之三部分
- 反映理論值與實際值之差異
 - 期刊數和文獻量統計不足，未能代表全部文獻之故。



理論與實際存在差異的原因

- 相關期刊數和文獻數統計不全所導致
- 選用布萊德福經驗法則，學科界限必須分明
 - 本應屬資訊科學，也有人會將其歸類於計算機科學，以致統計上有所誤差
- 應用布萊德福經驗法則，會借助於書目、索引、文摘之類的工具
 - 此類工具之編輯人員必須對原始文獻選擇和加工，可能改變原本的面目。
- 實際統計各種數據時，遇到的干擾
 - 期刊的更名、停刊、文獻的重覆發表等

布萊德福經驗法則的理論發展

- 布氏經驗法則的發展過程
- 威克瑞對布氏經驗法則的推論
- 萊古勒對布氏經驗法則的描述
- 布魯克斯對布定氏律的描述
- 斯馬里科夫的統一方程
- 布氏分布理論及發展趨勢

布氏經驗法則的發展過程

- 創立階段
 - 1934年，「特定主題的資訊源」此論文發表為起點
 - 1934~1948，並沒有引起學術界注意
 - 1948年，「文獻學」的出版，使其理論開始傳播
- 理論研究階段（60年代）
 - “布萊德福”熱-相關研究空前活躍
 - 形成兩個學派-區域派和圖像派
- 全面發展階段（60年代以後）
 - 由純粹理論研究向應用領域過渡
 - 理論與應用並舉之全面發展情形

威克瑞的兩個推論

- 布氏經驗法則不只侷限於劃分三個區域，其證明可將等級排列的期刊劃為任意多區

- 圖象表達式的推論（威氏公式）

$$n_1 : n_{1-2} : n_{1-3} : \dots : n_{1-m} = 1 : V : V^2 : \dots : V^{m-1}$$

$n_{1-k}(k=2,3,\dots,m)$ ---- 第1區到第k區的期刊累積量

m -----劃分的區域數

V -----分散係數（威氏係數）

- 布氏公式的推廣形式（布氏公式）

$$n_1 : n_2 : n_3 : \dots : n_m = 1 : a : a^2 : \dots : a^{m-1}$$

威式公式的證明

- 威式公式的正確性
 - 從圖形出發，利用相似三角形等幾何知識證明
- 說明了圖像描述法的正確性

威氏公式和布氏公式的比較

- 威氏公式和布氏公式不可同時並存
 - 但兩者皆可近似地作為研究期刊分散規律的計算公式
- 兩個公式之間的理論誤差趨於一致的條件
 - V (威氏係數) 充分大時
 - 相關論文高度集中於核心區少數幾種期刊，其他各區的平均載文率相差亦大

萊古勒對布氏經驗法則的描述

- 發展了區域描述方法
 - 從布氏經驗法則的文字敘述 (區域描述) 出發，推導出按等級排列的期刊中論文分布的規範化公式。
- 布萊德福累積分布函數的推導
 - $$F(x) = \frac{\ln(1+\beta x)}{\ln(1+\beta)}, \quad X(f) = \frac{b^f - 1}{b - 1}$$
$$b = a^2, a \text{ 是布萊德福乘數}$$
$$\beta = b - 1$$
 - x 表用來定義全部期刊中含某一指定數量相關論文的最少期刊與期刊總數的比例。
 - $F(x)$ 表示在 x 的比例下，期刊論文數與期刊論文總數的比例
 - 例如：要得到80%的論文，將 $F(x)=0.8$ 和相應的 β 值代入，反求出 x ，即所需期刊的最少數量。

布魯克斯對布定氏律的描述

- 發展了圖像描述法
 - 利用數學公式描述了布萊德福的經驗法則
 - 創造性地用兩個數學表達式描述布氏經驗法則
 - 圖像的曲線部分： $R(n) = \alpha n^\beta, 1 \leq n < c$
 - 圖像的直線部分： $R(n) = k \ln(n/s), c \leq n < N$
 - $R(n)$ - 對應於 n 的相關論文累積數
 - n - 期刊等級排列的序號、 α - 第一級期刊中相關論文數 $R(1)$
 - c - 核心區的期刊數、 N - 等級排列的期刊總數
 - k - 參數。分數曲線中直線部分的斜率，當 n 足夠大時， $k=N$
 - s - 參數。其數值等於圖形直線部分反向延伸與橫軸交點之 n 值
- 發現
 - 學科範圍擴大， s 值也增大，且與學科發展階段有關
 - s 可作為比較學科幅度和成熟程度時的參考
 - c 值與 s 值相關

斯馬里科夫的統一方程

- 提出一個統一方程來代替布魯克斯的兩個公式
 - $R(n) = k \log(q_1 n + q_2 e^{-\beta n})$
 - 利用此方程所繪的分布曲線與實際數據吻合度高
 - 可進一步地證明威氏公式的正確性。

布氏分布理論的基本內容

- 一個經驗法則、十來個經驗公式
- 一個經驗法則
 - 以“布萊德福分散經驗法則”為核心
- 十來個經驗公式
 - 威克瑞
 - 布魯克斯
 - 肯德爾之“布萊德福-齊普夫經驗法則”
 - 萊古勒的規範化公式
 - 斯馬里科夫的統一方程
 - ...

布氏分布理論的發展趨勢

- 主要之研究工作
 - 進行具體統計，驗證布氏經驗法則，並實際應用
 - 尋求普遍而精確的經驗分布公式和理論解釋，並取得了較大的發展
- 當前研究方向和亟待解決之問題
 - 運用大量之統計數據、嚴密的數學方法，確立或尋求更為精確的規範化之數學模型
 - 深入研究布氏分布的機制，尋求科學統一的理論解釋
 - 努力結合實際，開展應用研究，提高圖書資訊服務的效率

布萊德福經驗法則的應用

- 布氏經驗法則應用的基本方法
- 布氏經驗法則應用的主要領域
- 布氏經驗法則應用的條件與侷限

布萊德福應用實例

- 假設N期刊總數為400, R(N)文獻累積總數為2000
- $R(n) = N \ln(n/s)$, $R(N) = N \ln(N/s)$
- $p = R(n)/R(N)$

- $N=400$, $R(N)=2000$, $\ln(s)=1$, $p=(\ln(n)-1)/5$

涵蓋文獻數量 (%)	需要文獻來源數量	文獻來源 (%)
100	400	100
80	148	37
60	55	14
40	20	5
20	7	2

布氏經驗法則應用的基本方法

- 布氏經驗法則之主要應用
 - 等級排列技術和分析方法之應用
- 應用之三大步驟
 - 選用統計工具並獲取原始數據
 - 依據研究對象和目的
 - 等級排列統計數據
 - 按照期刊的載文率大小順序排列，是最重要的關鍵步驟
 - 分析統計資料，得出統計結果
 - 區域分析法 (Zone Analysis)
 - 圖像分析法 (Graphical Analysis)

布氏經驗法則應用的主要領域

- 指導期刊訂購工作
 - 確定“核心期刊”，為期刊選訂提供依據
- 用於文獻檢索
 - 預計完全檢索n種期刊的論文總數
 - 評價文獻檢索的效率
- 考察專著的分布
 - 分析各個出版社關於某學科或專業著作之出版情況，確定此學科的“核心出版社”，以指導圖書館的採購工作。
- 動態館藏的維護
 - 可分析流通期刊之最小核心，以及核心讀者所感興趣的學科之最小核心，準確地為收藏服務，使館藏保持為一種有序狀態

布氏經驗法則應用的主要領域 (續)

- 檢索工具完整性的測定
 - 用來確定某一覆蓋度的文摘索引至少要摘索多少種資訊源
 - 評價某一學科之檢索工具 (文摘、索引、目錄) 的完整性，為這些工具之選擇和利用提供科學依據。
- 學科幅度的比較
 - 可確定不同學科之核心區和s值
 - 核心區出現的期刊數量可作為兩門學科重疊程度的依據
 - s值可用以比較學科之領域範圍和發展的成熟程度
- 指導讀者利用期刊
 - 以“核心期刊”的概念指導讀者利用期刊

布氏經驗法則應用的條件與局限

- 論文的學科、專業領域或課題範圍應當清楚地劃定。
- 被分析的相關學科、領域或課題的期刊清單，以及對這些期刊中刊域的相關論文之統計應當是充分的。
- 被分析的期刊之時間應當清楚限定，以保證有關文獻數據統計之一致性。



結語與討論
