

書目計量學

Lecture 03 -- 學術文獻的成長

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

- 文獻資訊流的特性及成長規律研究的意義
- 學術知識量的成長與學術文獻的成長
- 學術文獻的指數成長律
- 學術文獻的邏輯成長律
- 學術文獻成長的其他數學模型
- 學術文獻成長機理的分析
- 學術文獻成長規律的應用

文獻資訊流的特性

- 文獻資訊流（文獻流）
 - 文獻所含資訊的匯流，是具有一系列主題特徵的學術文獻之集合。
- 文獻資訊流的特性
 - 靜態特性
 - 在一定時間內學術文獻在空間的分布性質，如集中—離散分佈、引文分佈、序性分佈、主題相關分佈等的規律性。
 - 動態特性
 - 學術文獻隨時間的延續而成長和老化的性質
 - 學術文獻既成長又老化，老化之中不斷的成長；成長是文獻資訊流的主要趨勢
- 學術文獻成長
 - 文獻數量隨著時間的推延而成長的情況
 - 美國《化學文摘》連續發表100萬篇文摘所用的年數不斷地縮短，從早期（1907~1938）需32年的時間，至今日只需2年的時間。

文獻成長的影響及對策

- 文獻成長的影響
 - 影響資訊工作的效率和資訊事業的發長
 - 圖書館與資訊機構對館藏文獻的儲存和管理非常困難
 - 影響學術研究效率和學術事業的發展
 - 文獻數量愈來愈大，研究人員難以查全和查準，學術研究會有重複和浪費的情況
- 文獻成長的對策
 - 理論上，加強文獻規律之研究
 - 文獻成長規律、老化規律等研究
 - 技術上，採用計算機等的先進技術和設備處理和利用文獻資訊

文獻成長規律的研究及意義

- 文獻成長規律研究的意義
 - 大致地揭示學術發展的特點和規律
 - 預測文獻成長的趨勢，為學術資訊工作未來的發展提供決策依據
 - 有效管理文獻資訊
- 文獻成長規律研究的現狀
 - 理論研究
 - 如何建立數學模型及理論解釋，以進一步地揭示學術文獻的成長規律
 - 應用研究
 - 如何運用文獻成長規律指導實際資訊工作和資訊管理
 - 如何將文獻數量指標運用於度量知識，以揭示學術發展規律

學術知識量的成長趨勢

- 學術知識量倍增的周期愈來愈短
- 學術發展的加速趨勢
 - 學術知識量的指數成長律？

學術知識量的成長與學術文獻成長

- 學術文獻是知識的客觀紀錄
- 學術文獻的數量
 - 其變化直接反映了學術知識量的變化
 - 學術發展的重要標誌，揭示學術發展的某些特點和規律
- 學術文獻量與學術知識量，具有同步成長的趨勢，其成長規律亦有很大程度上的相似性
 - 學術文獻成長規律的發現可以為學術知識量成長規律的研究提供依據
 - 學術知識量成長規律的研究將有助於加深對文獻成長規律的認識

文獻成長的幾項思考

- 假設 $P(t)$ 代表 t 時間點的文獻總數
- 文獻是以一定的固定比率成長
 - $dP(t)/dt=b$
 - 文獻線性成長
- 文獻是以一定的相對比率成長
 - $dP(t)/dt=bP(t)$
 - 文獻持續成長，而無止盡
 - 發散
- 文獻的成長比率與當時文獻總數有關，而非總是以相同相對比率成長
 - $b=k(M-P(t))$, assume M is maximum of literature
 - $dP(t)/dt=k(M-P(t))P(t)$
 - 文獻雖持續成長，但愈來愈慢，終究達到一最大值
 - 收斂

文獻成長的幾項思考 (續)

- 文獻成長亦非僅是文獻成長之內在規律，與社會環境、學科領域、生產技術、科研經費等許多外在因素影響
- 文獻並非持續成長無止盡，亦非終究止於最大值
- 文獻成長有時符合一規律，有時符合另一規律，有時很難解釋

學術文獻的指數成長律

- 文獻量度指標與方法
- 文獻指數成長模型
- 文獻指數成長規律的分析

文獻量度指標與方法

- 文獻計量的指標
 - 絕對值指標
 - 表示文獻數量的多少，如：圖書、期刊、論文數量。
 - 相對值指標
 - 表示不同部分文獻的數量比例，某一部分文獻占全部文獻的比例、各類型文獻的比例、各語種文獻的比例等等。
- 衡量學術文獻成長之方法
 - 非累積量：每年新出版的文獻數量
 - 較少研究，但能直觀地反映出每年新文獻的變化趨勢。
 - 累積量：每年出版的文獻之累積總量
 - 一般研究，大多以累積量出發，較有可能趨於某種固定的規律，容易用一個較為準確的函數來描述，有利於定量分析。

普賴斯曲線文獻指數成長模型

- 以學術文獻量為縱軸，以歷史年代為橫軸，得到一曲線，呈現學術文獻“按指數成長的規律”。

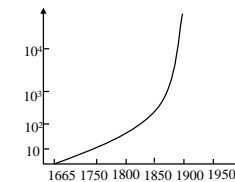
$$P(t) = ae^{bt} \quad (a > 0, b > 0)$$

t -- 時間，以年為單位

a - 條件常數，即統計的初始時刻
($t = 0$) 的文獻量

e - 自然對數的底

b - 時間常數，即持續成長率



- 不同學科的文獻成長速度不同

文獻指數成長規律的分析

■ 文獻指數成長律的正確性

將普賴斯指數方程， $P(t) = ae^{bt}$ ，求一階導數

得出曲線成長率： $\frac{dP(t)}{dt} = abe^{bt} = bP(t)$

相對成長率： $\frac{dP(t)}{dt} / P(t) = b$

因 $a > 0, b > 0$ ；則 $\frac{dP(t)}{dt} > 0, (0, \infty)$

- 數學意義：普賴斯成長函數在區間 $(0, \infty)$ 上是一單調遞增函數。

文獻指數成長律的侷限性

- 學術文獻並不總是按指數函數關係成長
 - 統計時間和文獻類型會有所影響
- 存在侷限性的原因
 - 沒有考慮文獻日益嚴重的老化因素
 - 在統計某年的學術期刊累積總數時，並無排除已停刊的期刊
 - 普賴斯的文獻指數成長律的結論只建立在學術期刊此種出版類型的文獻和某些知識領域的文獻在其統計的時間範圍內的成長特點的基礎上

學術文獻的邏輯成長律

- 文獻邏輯成長模型
- 文獻邏輯成長規律的分析
- 文獻邏輯成長模型的修正

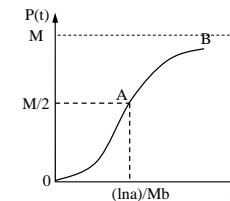
文獻邏輯成長模型

■ 文獻邏輯成長模型之數學表示式

$$P(t) = \frac{M}{1 + ae^{-Mbt}} \quad (b > 0)$$

$P(t)$ ----- t 年的文獻累積量
 M ----- 當 $t \rightarrow \infty$ 時文獻的累積量，即文獻累積量之最大值
 a, b ----- 參數

- 將上式對 t 求二階導數，並令二階導數為 0，求得此式之曲線的拐點 A 之坐標為 $((\ln a)/Mb, M/2)$
 - 當 $t < (\ln a)/Mb$ ，文獻急遽成長，成長率漸增。
 - 當 $t > (\ln a)/Mb$ ，文獻成長減慢，成長率是漸減的。



文獻邏輯成長模型 (續)

- 文獻邏輯成長模型之意義
 - 當 $P(t)=y < M$ 時，
 - 曲線成長率： $dy/dt = Mby$
 - 曲線的相對成長率： $\frac{dy}{dt} / y = Mb = const$
 - 學術文獻成長的初始階段，符合指數成長規律；當文獻量成長到最大值的一半 ($M/2$) 時，其成長率開始變小，最後逆緩成長，並趨近於一個極限值 $y=M$

文獻邏輯成長規律的分析

- 文獻邏輯成長律的正確性
 - 相關統計數據表明
 - 當學科處於誕生和發展時期：學術文獻呈指數成長
 - 學科的研究深入，進入相對成熟時期：成長率變小，曲線變得平緩。
- 文獻邏輯成長律的侷限性
 - 學術發展到一定階段時，學術文獻的成長率趨於0
$$P(t) = \frac{M}{1 + ae^{-Mbt}} \quad (b > 0)$$
當 $t \rightarrow \infty$ 時， $y \rightarrow M$ ， $dy/dt = by(M - y) \rightarrow 0$
 - 但科技文獻數量成長速度的減緩，並不意味著學術發展的速度會下降
 - 新的、更完善的傳播科技資訊的方法和手段出現，以補充或逐步取代現有傳統形式的科技文獻，學術研究和發展仍持續提昇。

文獻邏輯成長律的侷限性

- 邏輯曲線成長模型，部分符合指數曲線，但指數曲線成長模型已存在一定的侷限性。
- 指數曲線和邏輯曲線對科技文獻成長所作的預測，是依據預測學中的趨勢外推法
 - 學術文獻作為學術交流這一複雜系統中的子系統，成長規律受到許多方的影響
 - 利用系統論的觀點對其作系統分析，才可望獲得比較符合實際的結果

文獻邏輯成長模型的修正

- 弄清學術文獻成長的基本過程
 - 由指數式成長向邏輯式成長的轉變
 - 由邏輯式向直線式成長轉變
- 必須把學術交流系統作為一個整體來研究
 - 學術文獻體系是學術資訊交流系統的一個子系統，資訊交流系統中各子系統間的相互作用是非線性的。
- 抓住主要的決定性因素，忽略其他較為次要的因素影響。
 - 科研經費的成長、科技人員的增加
- 確定具體模型的目標和要求
 - 必須能克服指數曲線“發散”及邏輯曲線“有限”的困難
 - 符合學術文獻發展的實際過程和趨勢，能夠對文獻成長規律作較圓滿的解釋。

學術文獻成長的其他數學模型

- 線性成長模型
- 分級滑動指數模型
- 超越函數模型
- 舍-布成長模型

線性成長模型

- 線性成長模型的內容

$$P(t) = bt + a$$

$P(t)$ ----- t 年的文獻累積數
 b ----- 文獻的年成長率
 a ----- 當 $t=0$ 時的文獻數量

- 線性成長模型的分析
 - 1960-1972年全世界出版的圖書和小冊子數量成直線規律成長
 - 學術文獻線性成長模型適用於描述某些知識領域或某些類型的文獻成長
 - 學術文獻未來的發展將更多地傾向於直線模型

分級滑動指數模型

- 分級滑動指數模型的內容
- 分級滑動指數模型的分析

分級滑動指數模型

- 勒希爾 (Rescher)。出版物的數量成長與其質量有關，不同質量的成長速度是不相同的。
 - 文獻的質量等級指標 λ ，且 $0 \leq \lambda \leq 1$
 - $\lambda=1$ ：常規文獻（實際代表了全部文獻）
 - $\lambda=3/4$ ：有意義的文獻
 - $\lambda=1/2$ ：重要的文獻
 - $\lambda=1/4$ ：非常重要的文獻
 - $\lambda=0$ ：第一級（質等重要）的文獻
 - 如果時刻 t 的文獻總量為 $P(t)$ ，在 λ 級上的文獻量則為 $[P(t)]^\lambda$ ，各 λ 級的文獻成長方程式為

$$P(t)_{\lambda=1} = ae^{bt}$$

$$P(t)_{\lambda=3/4} = (ae^{bt})^{3/4}$$

$$P(t)_{\lambda=1/2} = (ae^{bt})^{1/2}$$

$$P(t)_{\lambda=1/4} = (ae^{bt})^{1/4}$$

$$P(t)_{\lambda=0} = \ln a + bt$$

分級滑動指數模型的意義

- $0 < \lambda \leq 1$ 時
 - 各級文獻仍按指數規律成長，但文獻質量提高時，隨著文獻之重要性程度的增大，成長速度減慢
- $\lambda = 0$ 時
 - 指數規律完全破壞，文獻數量呈線性函數成長

分級滑動指數模型的分析

- 不同質量的文獻，其成長速度不同，愈重要的文獻成長速度愈慢
- 數量較少的高質量論文總是伴隨著大量的一般性的論文而同時出現
- 很難用具體數據來驗證這一模型的正確性

超越函數模型

- 吉利亞列夫斯基，希萊德爾
- “吉-希成長模型”
 - 學術期刊論文數量的成長應考論期刊論文分散情形
 - 將某一學科或知識領域的期刊按布萊德福分布進行等級排列，則不同等級區域內期刊論文數的成長是不同的
 - 只是一種專門研究期刊論文數量成長的假說性之理論模型

舍-布成長模型

- 1978年，舍斯托帕爾，布爾曼
 - 舍-布成長模型的內容
 - 文獻的增加應是“減少的增加”；即相對成長率是隨著時間 t 的增加或文獻總量 M 的增大而減小
 - 文獻成長方程式（又稱為文獻資訊流的“總模型”）
- $$dP(t)/dt = q(t)P(t) \quad \begin{array}{l} P(t) \text{ ----- 文獻的累積總量} \\ q(t) \text{ ----- 文獻的相對成長率函數} \end{array}$$
- 舍-布成長模型的分析
 - 從 $q(t)$ 此一變量出發來研究文獻的成長規律，這是對以往研究的重要修正
 - 作為文獻資訊流的總模型是不夠全面的
 - 包括了文獻指數成長模型、也包括了線性成長模型和滑動指數模型
 - 但沒有包括邏輯曲線的成長模式

學術文獻成長機理的分析

- 學術文獻數量成長的原因
- 學術文獻成長規律的解釋

學術文獻數量成長的原因

- 科研經費和科技人員數量的激增
- 專業範圍的擴大和細分化
- 學科之間相互滲透
- 學術技術的國際化
- 研究的合作化和組織化
- 研究的週期較短、產生成果的速度加快
- 通訊、出版技術的改進和資訊工作的加強

學術發展規律

- 托爾斯·S·龐恩所提的學術發展模式
 - 原始科學→常規科學
 - 科學數量不大，但成長快速成指數規律成長，新思想和新成果的最重要論文出現時期
 - 常規科學→另一個常規科學
 - 科學已處於成熟和穩定發展時期，學術文獻數量龐大，但成長已減慢成邏輯由線模式；科學革命即將到來，學科範圍將發生變動
- Menard認為一門學科的文獻成長率隨不同時期而變化
 - 穩定界：直線型成長；學科處於誕生時期
 - 成長界：以較快的指數型速率成長；學科處於發展時期
 - 循環界：穩定和成長交替發生；學科處於成熟時期

社會環境條件

- 受到科學所處的環境條件，包括政治、經濟、文化、教育等社會條件等影響
- 受到載體技術、出版技術、電子計算機和現代通訊技術等眾多因素影響



學術文獻成長規律的應用

- 在科學和科技史研究中的應用
 - 用來判斷和預測知識的成長狀況，繼而探索整個學術的發展規律
- 在資訊研究中的應用
 - 作為研究人員和資訊工作人員進行資訊分析研究，掌握科技發展動態，進行科學預測的可靠方法
- 在圖書資訊管理中的應用
 - 圖書館的管理，確定經費的合理分配、資料搜集的原則、館藏增加的策略、貯存空間擴大的措施



結語與討論