Pauca sed Matura

De-Xing Guan (官德星, 國立台北大學經濟學系)

January 2021

我從小就很崇拜數學家 Gauss, 據說他不到十歲就發明從 1 加到 100 的公式, 這也是我常跟同學們分享的故事, 而他量少質精 (*pauca sed matura*) 的座右銘,[1] 更是令我心嚮往之。 除了質量的取捨之外, 他對常態分配的研究也對許多學科產生重大影響, 譬如財務和總體經濟學就常用到這個機率分配。 其實還有許多經濟理論和質量取捨或常態分配有密切關係, 然而有趣的是此二者都和 Coase 提出的交易成本 (*transaction cost*) 有關, 這使得我們可以透過交易成本的概念來重新探討相關經濟議題。 至於連結 Gauss/Coase 的橋樑則是資訊理論之父 Shannon 關於訊息傳遞的想法, 一個串聯機率、資訊和交易成本的有趣想法!

量少質精

Gauss 認為不成熟的想法不值得發表, 所以他堅守量少質精的原則, 然而質量真的無法兼顧嗎? 質量有可能與時俱進或同時下滑嗎? 還是二者是獨立的呢? 在資源有限下, 質量的取捨是常態, 因為大量生產時品質較難維持, 然而產業的創新仰賴品質的提升, 否則便容易被成本更低的廠商取代, 因此廠商也有升級的壓力。 創新需要研發, 所費不貲的研發成本是廠商能否升級的關鍵, 而保護智慧財產權和維持公司組織順利運作也是耗時費力的工作。 在有交易成本的情況下, 若提高品質所帶來的產值上升大於因此而產生的交易成本, 則公司獲利增加, 此時質與量 (公司產值) 與時俱進; 反之則無法兼顧, 甚至有可能同時下滑。

質量的取捨取決於交易成本的大小, 但是交易成本有好有壞, 好的交易成本 (譬如維護產權的成本) 會使質量俱增, 而壞的交易成本 (譬如意識形態的衝突) 則會使質量同時縮水。 交易成本和原始生產成本 (*prime cost*) 一樣, 都是生產要素用於其它地方所能得到的最大產值,[2] 只不過這些要素是用於和交易相關的行為, 而不是用於生產。 所以我們不能只從交易成本的增減來探討質量取捨的問題, 交易成本的結構也是很重要的。 這表示我們不能先驗上認為提高品質就

[1] Carl F. Gauss 用的是拉丁文, 直譯是稀少但成熟 (few, but ripe) 的意思。
[2] Ronald H. Coase, *The Firm, the Market, and the Law*, University of Chicago Press, 1988, p. 80。

一定是好的，大量生產就不可取，因為有時候先用數量搶占市場，再慢慢提升品質也是有效的競爭手段；反之，數大不一定就是美，盲目追求大數據 (*big data*) 不必然會使產銷更有效率，因為數據多，雜訊或噪音 (*noise*) 也多,[3] 而許多雜訊會產生不好的交易成本，使經營效率降低。

由於真實世界永遠存在交易成本，因此質和量並沒有必然的關係，而因為蒐集資訊成本是交易成本中最重要的一環,[4] 所以探討資訊如何形成、傳遞和使用，一直都是經濟學的核心議題。 譬如 Hayek 早期就以研究這些問題著稱，他認為價格機能是用來傳遞眾多消費者和生產者個別資訊最有效的途徑,[5] 而由於蒐集和傳遞資訊需要成本，因此沒有一個人或廠商擁有全部的資訊，換句話說，假設資訊完全的完全競爭市場在真實世界是不可能出現的。 然而大多數的經濟模型都假設完全競爭，因此在完全競爭下同時出現資訊不完全的唯一可能，就是假設那些無法取得的資訊或噪音是隨機變數，如此不僅可以規避蒐集資訊成本 (因為完全競爭模型假設交易成本為零)，還可以考慮不完全資訊的影響，可說是一舉兩得，而常態分配是最常被使用的機率分配，因此在經濟模型中出現常態分配的假設並不令人意外。 問題是：常態分配和完全競爭的假設真的沒有矛盾嗎?

<u>常態分配</u>

由法國數學家 de Moivre 利用巴斯卡三角形 (*Pascal's triangle*) 發現的機率分配,[6] 後來被 Gauss 推廣成所謂的常態或高斯分配 (Gaussian distribution)，這個分配在 Laplace 的中央極限定理 (*central limit theorem*) 支持下，更成為眾多機率分配的基本代表款，而廣為統計學家和經濟學家使用。 這個分配有兩個重要假設：樣本數愈多愈好，而且彼此獨立,[7] 根據中央極限定理和 Jacob Bernoulli 提出的大數法則 (*law of large numbers*)，當樣本數不斷增加時，任何期望值和變異數固定的機率分配都會收斂到對稱的常態分配。 所以當經濟學家在使用常態分配時，他必須假設市場的人數或廠商夠多，而且這些消費者的偏好和生產者的技術都必須彼此獨立。 這個通常稱為相同且獨立分配 (identical and independent distribution) 的假設其實很嚴格，因為除了擲骰子之外很難找到其它實例。 不過也正因如此，它反而和完全競爭的假設相容，一點也不矛盾!

[3] Nassiam Nicholas Taleb, *Antifragile*, Random House, 2012, p. 418 (中譯：<u>反脆弱</u>，大塊文化, 2013, 頁 560)。 他認為: "雜訊呈現凸性，而資訊呈現凹性 (noise is convex and information is concave)", 所以雜訊增加的速度會超過資訊。 我們稍後會用 Shannon 的資訊理論來討論這個有趣的現象。

[4] 除了蒐集資訊成本外，另兩個交易成本是協商議價成本和執行合約成本，三者分別是交易前、交易時和交易後最主要的交易成本，細節請參考: Coase (1988), pp. 38-39。

[5] Friedrich Hayek, "The Use of Knowledge in Society," *American Economic Review*, 1945, 519-530。

[6] Peter L. Bernstein, *Against the Gods*, Wiley, 1996, p. 127 (中譯：<u>風險之書</u>，商周, 2019, 頁 135)。

[7] *Against the Gods*, p. 142 (<u>風險之書</u>，頁 149)。

我們可以用一個簡單的例子來說明常態分配和完全競爭是一致的。想像在股市中有人看多，有人看空，但是多空雙方的決策彼此獨立，互不相干。此時只要投資人數夠多，多方和空方的報價便會逐漸對稱於市場均價，譬如認為會漲 5% 的人數和認為會跌 5% 者相若，認為會漲停的人數也和認為會跌停者大致相當，於是平均而言，投資人的預期報酬是零，而這正是常態分配運作的結果。事實上遠在 1900 年，現代財務理論之父 Bachelier 就做了這樣的假設。[8]

最後我們來看為什麼根據常態分配運作的股市必然是完全競爭呢？首先在完全競爭市場中，買賣雙方都是價格接受者，這和常態分配下多空決策彼此獨立的假設一致。此外根據常態分配的對稱性，由於買賣雙方都知道看漲和看跌某個百分比的人數相同，因此沒有理由蒐集交易對手的資訊，也無須釋出假消息去放空或做多，因為大家都心知肚明而且人數和金額都相當，這表示買賣雙方都不必付出任何交易成本去傳遞和使用訊息，這和完全競爭下交易成本為零的假設是一致的。所以經濟學家喜歡常態分配不是沒有道理的，因為它恰好和主流經濟學常用的完全競爭模型相容，卻又不需考慮交易成本，難怪到處可見它的蹤影。然而經濟學家利用常態分配偷渡不完全訊息到完全競爭模型的想法，[9] 卻始終無法和真實世界的現象契合，因為儘管常態分配和完全競爭相容，但都與現實相距甚遠，這使得模型的解釋能力大為降低。最明顯的例子莫過於關於菲利浦曲線 (*Phillips curve*) 以及效率市場假說 (*efficient market hypothesis*) 的討論。我們先從第一個例子說起。

島嶼寓言

1969 年初有十四位經濟學家在費城舉辦了一個有趣的研討會，我說有趣是因為參與者的論文不是即將發表，就是已在別處發表，而會議召集人 Phelps 還規定每個與會者都只能評論別人的論文，不用報告自己的文章。至於會議的主題則是關於菲利浦曲線，也就是通膨和失業的相關議題，並且將重點放在不完全資訊模型。[10] 出席者後來獲得諾貝爾經濟學獎的有 Lucas, Phelps, Mortensen 等人；此外，張五常的恩師 Alchian 也在受邀之列，但未出席，而他也是唯一從成本角度來討論不完全訊息的受邀者。身為召集人的 Phelps 也在會中提出一個概念，

---

[8] "*The mathematical expectation of the speculator is zero*." Louis Bachelier, *Theory of Speculation*, Princeton University Press, 2006, p. 28。

[9] 即使是現在逐漸成為主流的獨佔性競爭模型也有同樣的問題，因為它也假設資訊完全。不過由於它放棄了價格接受者的假設，所以不太需要資訊不完全的設定，就可以解釋一些現象，但卻有寇斯猜想 (*Coase Conjecture*) 的問題，所以仍然不是一個好的理論模型。

[10] Edmund S. Phelps, "Introduction," in *Microeconomic Foundations of Employment and Inflation Theory*, Norton, 1970, pp. 1-23。不過 Phelps 將與會者提出的模型統稱為失衡 (disequilibrium)，而不是完全競爭模型，顯然他也不認為完全競爭和資訊不完全可以相容。

希望能成為不完全訊息模型的基礎，他稱此概念為島嶼寓言 (*island parable*),[11] 後來 Lucas 據此建立他著名的理性預期景氣循環模型,[12] 成為主流的總體理論架構。 島嶼寓言的基本概念很簡單：想像所有生產者分散在許多島嶼上，各島彼此獨立，互不往來，資訊也無法在島間流通，每個島的生產者都只能在自己的島上交易，但消費者可以跨島消費。

商品和資訊無法跨島移動是因為移動需要成本，也就是交易成本，可是島嶼寓言的目的是要建立一個不完全訊息和完全競爭相容的模型，以便解釋菲利浦曲線，但不能考慮交易成本，否則均衡就不會是完全競爭。 Lucas 的理性預期模型仍然假設完全競爭，但將資訊無法跨島移動視為生產者只能看到雜訊或噪音，再據此去推估需要的訊息，於是內生的交易成本變成外生的隨機變數，且為常態分配。 如前所述，常態分配和交易成本為零的假設是一致的，因此在 Lucas 的模型中並未出現交易成本，而短期貨幣不中立 (或菲利浦曲線為負斜率) 來自生產者資訊不完全，無法區別噪音導致的物價上漲是只來自本島，還是各島齊漲。

然而 Lucas 的解釋有個盲點，那就是即使像印尼這樣的多島國家，我們也很難想像商品和資訊無法跨島移動，否則受中央銀行影響的一般物價變動要如何讓所有百姓知道呢? 如果是透過消費者跨島消費來傳遞，那麼各島生產者為何不和這些消費者勾結，付費請他們散播不實消息來做多或放空呢? 此外, Lucas 也承認如果沒有島嶼寓言，亦即如果商品和資訊可以自由跨島移動，那麼貨幣不管長短期都是中立的，而菲利浦曲線也將是垂直線，因為即使是未預期到的貨幣波動也不會對商品的總合供給產生任何影響。[13] 這使得貨幣景氣循環必須仰賴實質面的波動才能產生作用,[14] 貨幣數量說 (*quantity theory of money*) 的支持者應該不會喜歡這個結果吧!

如果真實世界的廠商不是在各自獨立的島上生產和販售，人們的預期也不是理性的，不知道也不會計算常態分配的期望值和變異數，那麼基於島嶼寓言的理性預期模型對實際現象的解釋力就會降低。 如果放棄常態分配的假設，那麼交易成本就不能忽略，可是資訊、噪音和交易成本究竟有什麼關係? 風險和不確定性該如何度量? 而除了島嶼寓言，還有什麼方法可以討論不完全資訊的影響呢? 這些問題都不容易回答，但都必須回答，否則我們就只能一直做出 Coase 所謂黑板經濟學 (*blackboard economics*) 的假設而不可自拔。[15] 我也不知道答案究竟是什麼? 但是我知道這些問題的解答一定和 Shannon 的資訊理論脫不了關係!

---

[11] Phelps (1970), p. 7。

[12] Robert E. Lucas, Jr., *Studies in Business-Cycle Theory*, MIT Press, 1981。

[13] Lucas (1981), p. 7。

[14] 此處實質面的波動來自生產者跨島的隨機分配。 其實這也預告日後 Lucas 會轉向實質景氣循環 (*real business cycle*) 的解釋，因為單純的貨幣波動不足以解釋實際的景氣波動現象。

[15] Coase (1988), p. 28。

資訊與噪音

經濟學之父 Smith 在國富論一開始便提到製針工廠 (pin factory), 這可能和因改良蒸汽機聞名的 Watt 曾經是他任教的格拉斯哥大學的技師有關。[16] 蒸汽機的廣泛應用除了對經濟學有所啟發之外, 也直接催生了熱力學 (thermodynamics) 和相關的統計力學 (statistical mechanics), 而後者是將機率理論應用於熱力學。[17] 熱力學和統計力學的核心概念是所謂的熵 (*entropy*), 它是用來度量分子受熱時的無序 (disorder) 狀態, 當熵愈大時表示分子互動愈頻繁也愈混亂。 無序代表分子的自由度提高, 換句話說它的運動方式也就愈不可預測, 也就是不確定性愈高的意思。 如果以人們的行為來類比, 當一個人的選擇愈自由或選項愈多時, 我們就愈難預測他的選擇, 或可說其行為的不確定性提高了。[18]

Shannon 在去麻省理工教書之前, 曾在貝爾實驗室工作, 負責研究訊息如何傳遞, 當時電話還不普遍, 訊息多半利用電報來傳送, 然而電報線傳輸容量有限, 因此如何壓縮訊息, 用最有效的方式傳訊, 就成為研究的焦點。 Shannon 想到可以將訊息用 0/1 兩位數的方式數位化, 這是因為我們可以令電流通過時訊號為 1, 沒通過時訊號為 0, 於是所有訊號都可用一連串的 0/1 表示, 而兩個數位 (binary digit) 若保留頭尾, 便成為一個叫做位元 (bit) 的新字, 也就是資訊流量的單位。 然而若要有效傳遞訊息, 必先定義什麼是資訊?[19] 這個看起來無聊的問題, 其實非常重要, 這就是我常跟學生說的: 問題比答案重要, 而能問出好問題更重要!

為什麼 Shannon 問的問題重要? 因為不同的資訊有不同的傳遞和接收方式, 如果資訊的數量無法度量, 就不可能知道什麼是有效的傳收方式了。 譬如兩個人當面說話沒有字數限制, 可以暢所欲言, 但若是用電報聯繫, 就必須精簡才行, 因為電報線容量有限, 傳輸成本自然比面談高出許多。 因此若考慮傳遞資訊的交易成本, 顯然資訊也必須遵從 Gauss 量少質精的原則: 要用最少的容量表達最多的訊息! 當資訊傳遞成本高時, 就得廢話少說, 直接講重點。

---

[16] Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, 5th edition, 1789; Modern Library, 1994, Ch. 1。 Smith 是否認識 James Watt 並有所交流, 我們不得而知, 不過當時棉紡織業在英國開始興盛則是不爭的事實。

[17] Ian Stewart, *Seventeen Equations that Changed the World*, Profile Books, 2012 (中譯: 改變世界的 17 個方程式, 商周出版, 2013), Ch. 12。

[18] 熵的操作型定義來自統計力學之父 Ludwig Boltzmann, 而著名的熱力學第二定律 (*second law of thermodynamics*) 則認為熵不斷增加使能量不斷消耗是一個幾乎無法逆轉的現象, 除非出現想像中的馬克士威精靈 (*Maxwell's demon*), 或是產生所謂的負熵 (*negative entropy*), 才有可能逆轉熵極大化的現象, 細節請參考: Erwin Schrödinger, *What Is Life?*, Cambridge University Press, 1944 (中譯: 生命是什麼?, 貓頭鷹出版, 2016), Ch. 6。

[19] 有關資訊理論發展的簡介, 請參考: William Poundstone, *Fortune's Formula*, Hill and Wang, 2005 (中譯: 天才數學家的秘密賭局, 平安文化, 2008)。

然而無論如何精簡，資訊數量的減少和資訊品質的提升都仍會受到雜訊或噪音的限制，除了傳遞媒介本身的折舊或毀損之外，有的噪音是傳訊和收訊者有意無意之間產生的，譬如故意散播不實訊息以影響股市、投票，或是訊息的誤植、誤信、誤傳等不當行為，無論是有心還是無意，這些行為都會使訊息變得模稜兩可，從而降低資訊傳遞的效率。 這表示傳出的訊息和收到的訊息不會完全相同，中間含糊其辭 (*equivocation*) 的部分，可用來度量訊息的平均模糊程度 (average ambiguity)。[20]

如前所述，無序代表選擇愈自由，也愈難預測。 以擲骰子為例，若每一面出現的機率都是 1/6，我們無法預測接著會擲出哪一點，但如果其中一點被人灌鉛，使其出現機率增為 1/3，此時知道內情的人必然會將加倍的賭注押在該點數上。所以若骰子是公正的，結果最難預測，此時會有最大的熵值 (*maximum entropy*)；而若骰子不公正，則結果較易預測，此時熵值也會較低。 由於擲骰子是我們能找到的最接近常態分配的實際現象，因此我們大概可以猜到：相對所有期望值和變異數都是有限值的機率分配而言，常態分配會有最大熵值。[21]

Shannon 利用無序和不確定的關係推導出一個度量資訊的熵值，若以最簡單的擲銅板為例，這個熵值 H 可以寫成： $H = -(p\log p + q\log q)$，其中 p 和 q 分別為出現正面和反面的機率，log 為以 2 為基底的對數，因為資訊是用 0/1 這兩位數的序列所組成。[22] 從這定義可以清楚看到當銅板是公正的時候，正反面出現的機率都是 1/2，此時 H 會達到最大，而當擲銅板次數不斷增加時，其機率分配也會收斂到常態分配，此時相對其它分配會有最大的熵值。

公正的骰子或銅板傳遞出來的資訊，幾乎沒有含糊其辭的部分，因為每一面出現的機率永遠都相等，沒有散播不實訊息或是上下其手的誘因，所以賭徒都只能純粹碰運氣！此時交易成本趨近於零，而機率分配也會趨近常態，這再一次證明我們在前面的猜測是正確的，亦即常態分配和交易成本為零的假設是一致的。如果我們用 Shannon 的資訊熵值來看，這表示交易成本為零和最大熵值也是一致的，而交易成本為零又和完全競爭相容，所以最後我們會得到一個有趣的結果，那就是：完全競爭、常態分配、交易成本為零、最大熵值這四個概念都是一致的。

當然我們知道在真實世界這四件事都很難出現，因為那是一個長期無所爭的寂靜世界，一個 Demsetz 戲稱的涅槃 (*nirvana*) 世界。[23] 所以如果我們回到真

---

[20] Claude E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 1948, pp. 379-423 and 623-656, 有關“含糊其辭”的討論在 p. 407。

[21] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, 2nd edition, Wiley, 2006, Ch. 12。

[22] Shannon (1948), p. 394。

[23] Harold Demsetz, “Information and Efficiency: Another Viewpoint,” *Journal of Law and Economics*, 1969, pp. 1-22, 有關“涅槃”的討論在 p. 1。

實世界，那麼市場不再是完全競爭，常態分配不再是常態，交易成本也無所不在，而熵值則會受含糊其辭的限制而變小，成為限制下的極大化問題。 Shannon 對這些問題的回應是提出了非常重要的夏農定理 (*Shannon Theorem*)，他證明容量 (capacity) 或最大的資訊傳輸率 (information rate of transmission) 等於傳訊者的資訊熵值減掉含糊其辭的熵值，後者是考慮收訊者狀況的條件熵值 (*conditional entropy*)，亦即：$C = \max(H(x) - H(x/y))$，其中 x, y 分別代表傳訊者和收訊者。[24] 含糊其辭會使傳遞資訊產生交易成本，降低資訊傳輸率，就像管線堵塞一樣，無法貨暢其流；反之，如果語意清晰，意簡言賅，則管線通暢，頻寬加大，就像 5G 上路一樣，資訊傳輸率大為提升。 所以當含糊其辭代表的條件熵值為零時，資訊傳輸率達到極大，並等於最大熵值；而當條件熵值等於傳訊者的資訊熵值時，資訊傳輸率為零，此時雜訊或噪音太大，資訊傳遞完全失靈。

Shannon 的資訊理論一開始和經濟學沒什麼關係，除了 von Neumann 建議他把資訊的定義取名為熵之外,[25] 同校任教的 Samuelson 則並不認為他的理論可以應用於經濟學。 所以讓資訊理論和經濟學產生連結，並將夏農定理發揚光大的，不是經濟學家，而是 Shannon 在貝爾實驗室的同事 Kelly，一位英年早逝的物理學家。 他因為喜歡賭馬而將夏農定理用來研究如何從投資中獲取最大利益，但是和那些相信隨機漫步 (*random walk*) 的效率市場支持者不同的是：他不認為個人無法打敗市場，尤其是當你擁有一些私人管道 (*private wire*) 的時候![26]

市場與效率

Bachelier 是將隨機漫步假設用於股市分析的始祖，而且如前所述，他認為投資人的預期報酬是零，這表示股價在理論上可以為負值，可是真實世界的股市並不容許負值，因此如果隨機漫步和相關的效率市場假說是對的，那麼當股價的預測值變成負值時，就會和實際股價產生較大的落差。[27] 隨機漫步的殘差項通常假設為相同且獨立的常態分配，且期望值為零，因此以它為基礎的效率市場假說就和島嶼寓言如出一轍，二者都隱含交易成本為零，因此每個人都沒有誘因蒐集和

---

[24] Shannon (1948), Theorem 11, p. 411。其實這篇文章有許多定理，但一般所謂的夏農定理指的是定理 11，而因熵是由機率組成，所以等式是幾乎確定 (almost surely) 相等，亦即誤差可以極小。
[25] 賽局理論之父 John von Neumann 認為 Shannon 對資訊的定義和 Boltzmann 對熵的原始定義很像，於是做了這個建議，但其實二者只是形式類似，實質意義並不相同。 Paul Samuelson 對 Shannon 的理論應用於經濟學一直不以為然，我只能說文人相輕，自古皆然吧!
[26] John L. Kelly, Jr., "A New Interpretation of Information Rate," *Bell System Technical Journal*, 1956, pp. 917-926。 私人管道可能是內線消息 (inside information)，也可能是憑本事取得的資訊優勢。
[27] 這個現象也發生在去年的輕原油期貨以及泰勒法則 (*Taylor rule*) 對次級房貸後的利率預測上，這是因為泰勒法則容許名目聯邦基金利率出現負值，但 Fed 不容許，於是其預測出現大幅落差。 關於負油價事件，請參考: https://web.ntpu.edu.tw/~guan/courses/NegativePrice.pdf

傳遞訊息，因為這些私人訊息平均而言都不會有任何作用，在大數法則運作下，小島的生產者和股市投資人都跟擲公正骰子的人一樣，只能純粹碰運氣。 可是如此一來，廠商和股市都沒有必要存在，而企業家和分析師也不會出現，這樣的效率市場可能離涅槃比離現實更近!

Shannon 的資訊理論如果對經濟學有任何啟示，那必然是來自資訊和市場效率的關係。 經濟學所謂的效率主要來自競爭，所以完全競爭是最有效率的均衡，而完全競爭假設完全資訊、自由進出和商品同質，這使得廠商無法透過產品差異化或是設下進出障礙來互相競爭，於是只剩下不完全訊息這個競爭手段，也就是比誰知道的多一點。 然而在常態分配的假設下，我知道比你多一點跟少一點的機率是一樣的，無論我怎麼累積私人資訊，平均而言都沒有幫助，就跟什麼都不做一樣，於是知識不再被累積，也不重要，反正最後都是用擲骰子決定。 涅槃本應在長遠的未來，但在完全競爭和常態分配的假設下，卻是每天的日常，這就難怪 Hayek 會說完全競爭其實是完全沒有競爭,[28] 因為每個人或廠商都沒有競爭手段可用，而且即使有，平均而言也沒有用。

如果我們放棄完全競爭和常態分配的假設，回到存在交易成本的真實世界，那麼我們要問的問題就不再是要如何達成完全競爭，因為完全競爭市場最有效率，而是要如何在不完全競爭市場下定義什麼是效率? Kelly 認為應該從成本最低著手，但是因為他無法定義一個適當的成本函數，只好討論如何在有私人管道下追求最大收益，因為他認為這是實際的賭徒或投資人唯一會問的問題。他和 Shannon 一樣認為當投資人有私人管道時，便具備了一些優勢 (*edge*)，他認為這就類似 Shannon 所謂的資訊傳輸率，但也同時會有一些無法控制的噪音，譬如跑馬當天的狀況，或是會影響股市的政經訊息等，所以當投資人在給定噪音下，將優勢發揮到極致時，就會達到最大的資訊傳輸率。

Kelly 利用夏農定理導出一個重要的凱利公式 (*Kelly formula*),[29] 而以此公式作為投資的判準被 Thorp 稱為凱利準則 (*Kelly criterion*),[30] 後來被一些投資人奉為圭臬，有趣的是 Samuelson 和效率市場的支持者始終排斥它，並認為它是一個謬誤 (fallacy)。[31] 無論誰對誰錯, Shannon/Kelly 告訴我們: 效率來自競爭，

---

[28] Friedrich Hayek, “The Meaning of Competition,” in *Individualism and Economic Order*, University of Chicago Press, 1948, pp. 92-106。

[29] 凱利公式最簡單的形式就是: *edge*/*odds*，也就是優勢除以賠率。 譬如你的賭本是 100，賠率是 5 倍，而根據私人管道你認為會贏的機率是 1/3，這表示你有 1/3 的機會得到 600 (100 本金加 500 彩金)，因此平均而言，來自私人管道資訊的價值是 200 (=600*1/3)，亦即私訊所帶來的報酬率或優勢是 1 (=(200-100)/100)，此時優勢除以賠率為 1/5，所以凱利準則告訴我們應該將 1/5 的資金拿去下注會得到最大的報酬率，參見: Poundstone (2005), pp. 72-73 (天才數學家的秘密賭局, 頁 77)。

[30] Edward O. Thorp, *A Man for All Markets*, Random House, 2017 (中譯: 他是賭神，更是股神，商周出版, 2018), Ch. 27。

[31] Poundstone (2005), p. 211 (天才數學家的秘密賭局, 頁 223)。

而競爭來自更多的選項和選擇方式，但噪音無可避免，它是競爭的一部分，因為沒有噪音，資訊傳遞就沒有差異性，也就喪失了效率，任何人要在市場獲利，不是因自己做對，就是靠別人犯錯，如果沒噪音，任何人都不會有資訊優勢或劣勢，所以噪音有弊也有利，噪音和交易成本一樣有好也有壞。 曾任美國財務學會會長，並提出選擇權定價公式的 Black，就認為金融市場交易的是噪音，可謂一針見血之論。[32]

當然噪音不能都來自內線消息或是官商勾結，這些是負面的噪音，只會擾亂交易，使市場效率降低；如果噪音反映的是不同的想法，或是為了保護產權而設定的編碼 (coding)，那麼反而會促進市場競爭，提高交易的效率。 來自私人管道的資訊有好有壞，但如果取得、傳遞和使用這些資訊都不會讓擁有或使用資訊者得到任何優勢，那麼這些資訊一開始就不會產生，而透過資訊連結而產生的知識也不會累積，經濟和社會的發展不是停滯，就是已經進入涅槃的極樂世界，後者是什麼我們不知道，但前者應該不是我們所樂意看見的。

如前所述，Kelly 在推導凱利公式時原本是要用追求成本最小為出發點，但是他不知道要如何設定成本函數，於是作罷，改為追求最大收益。[33] Kelly 會這樣想，是因為他知道追求最大資訊傳輸率，和追求最小含糊其辭的條件熵值是同一回事。[34] 含糊其辭類似經濟學中的交易成本，Shannon/Kelly 用條件熵值來表示，那麼經濟學家應該用什麼來量化交易成本呢? Stiglitz 是一個著名的例子，他曾經利用固定成本來代表取得資訊的交易成本，從而證明 Hayek (1945) 的想法是對的,[35] 然而交易成本不只是一個不隨交易量改變的固定成本，而是一個會隨交易方式而改變的變動成本，因此 Stiglitz 還是沒有抓到重點: Hayek 的想法其實和 Coase/Shannon/Kelly 比較接近，而不是 Samuelson/Stiglitz!

在這即將進入 5G 的時代，應用大數據的科技紛紛出籠，舉凡機器學習、深度學習、人工智慧等需要大量數據和快速運算的應用都和機率和資訊理論有關，譬如貝氏定理 (*Bayes Theorem*) 就是這些新科技運用不可或缺的工具。[36] 但是資訊的數量增加的同時，資訊的品質和市場效率也會跟著提升嗎? 我不知道答案是什麼，但我只能說資訊就像是中間投入的生產要素，它需要勞動和資本等原始投入要素去生產，而它本身又是最終產品的投入要素。 因此資訊產權的界定非常重要，因為權利的界定是市場交易的前提,[37] 而在討論資訊數量和品質問題之前，也應先將雜訊或噪音視為一個生產要素，而不是隨機變數或是固定成本啊!

---

32 Fischer Black, “Noise,” *Journal of Finance*, 1986, pp. 529-543。

33 Kelly (1956), p. 918。

34 Kelly (1956), p. 917。

35 Sanford J. Grossman and Joseph E. Stiglitz, “On the Impossibility of Informationally Efficient Markets,” *American Economic Review*, 1980, pp. 393-408。

36 William Poundstone, *The Doomsday Calculation*, Little, Brown Spark, 2019。

37 Coase (1988), p. 158。