數學百子櫃系列(二)

漫談數學學與教

新高中數學課程延伸部分

單元一

作者 韓藝詩、黃毅英、張家麟

教育局
課程發展處數學教育組

ISBN 978-988-8019-63-2

目錄

前言

為配合香港數學教育的發展，並向教師提供更多參考資料，課程發展處數學教育組於 2007 年開始邀請大學學者及資深老師撰寫專文，並蒐集及整理講座資料，輯錄成《數學百子櫃系列》。本書《漫談數學學與教 — 新高中數學課程延伸部分單元一》是這個系列的其中一冊，作者黃毅英教授、張家麟博士和韓藝詩女士對中學數學教學素有研究，本書除談及高中數學課程的學科知識外，對學科教學知識、學習難點等，都有精闢的見解。本書不僅可供教師參考，亦可作為學生讀物。作者撰文期間，高中數學課程仍在修訂，本書內容或與課程最後定稿偶有出入，祈請讀者留意。此外，本書只屬作者個人觀點，並不代表教育局的意見。

本系列能夠出版，實在是各方教育工作者共同努力的成果。在此，謹向提供資料、撰寫文章的老師、學者，以及所有為本書勞心勞力的朋友，致以衷心的感謝。

如有任何意見或建議，歡迎致函：
九龍油麻地彌敦道 405 號九龍政府合署 4 樓
教育局課程發展處
總課程發展主任（數學）收
(傳真：3426 9265 電郵：ccdoma@edb.gov.hk)

教育局課程發展處
數學教育組

作者簡介

黃毅英，文學學士、哲學碩士、教育證書、哲學博士（香港大學），文科教育碩士（香港中文大學），現任香港中文大學課程與教學學系教授。於境內外學報發表學術論文二百餘篇。2001 年獲香港研究資助局重點專案資助（Competitive Earmarked Grant）、2005 年獲學院優秀教學獎、2006 年第三屆全國教育科學研究優秀成果獎三等獎、2008 年獲香港中文大學研究卓越獎。編著有《邁向大眾數學的數學教育》、《數學教育實地觀察》、《數學教育實地再觀察》、《香港近半世紀漫漫「數教路」：從新數學談起》、《華人如何學數學》（與范良火、蔡金法、李士錡合編）、《迎接新世紀：重新檢視香港數學教育 — 蕭文強教授榮休文集》、《香港近半世紀漫漫「小學數教路」：現代化、本土化、普及化、規範化與專業化》（與鄧國俊、霍秉坤、黃家樂、顏明仁合寫）、《變式教學課程設計原理：數學課程改革的可能出路》（與林智中、孫旭花合寫）等。香港數學教育學會創會會長，現爲上海師範大學小學教育研究所客座研究員、天津《數學教育學報》及韓國《數學教育研究學報》編委。

張家麟，學數於香港中文大學數學系，先後獲學士、碩士及博士學位。研究興趣為非綫性偏微分方程。曾任職中學教師、香港教育學院及香港中文大學數學系導師，2005 年獲中文大學理學院優秀教學獎。2006 年 7 月任香港教育學院助理教授至今，對數學解難，以及幾何的學與教至感興趣。

韓藝詩，香港科技大學獲得理學士（數學）和哲學碩士（數學）學位，香港浸會大學取得學位教師教育文憑，現於香港中文大學修讀教育碩士課程。曾於香港教育學院擔任專任導師，亦曾於中學任教數學。

1. 緒論

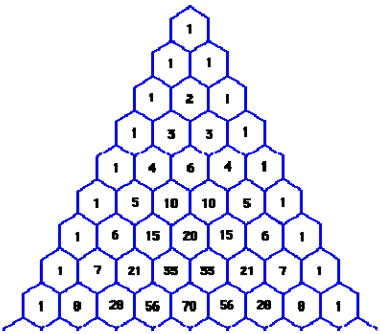
新高中數學課程單元一與數學及統計科(高級補充程度)內容相若，並循著數統的脈絡，由三個領域組成。第一個領域「基礎知識」的內容是為了學習第二個領域「微積分」及第三個領域「統計」而鋪路。例如，有了二項展式就可介紹二項分佈；有了 e 就可介紹正態分佈。第二個領域「微積分」基本上是單元二的子集。當然，相對單元二而言，它較強調運算的技巧和在統計方面的應用。

在編寫此小冊子時，筆者們曾考慮刪除「微積分」部分。因為這部分與單元二有重複的地方，雖然深度、廣度和重點不同，老師們為了有更詳盡和深入理解，想必直接翻閱單元二的小冊子，所以在編輯之初是沒有「微積分」部分。但參考多方友好的意見後，為了方便老師閱讀，所以部分單元二的內容會直接抽取並置於單元一的相關課題內，故發現單元一與單元二部分內容相同是正常的。

這本小冊子並不是一部數學書，想查閱對課題更深入的數學內容與証明，恐怕要讓老師們失望了。這亦並非一部教學法或提供教學範例的書籍，想把此書當作教學指引的老師們也許亦得不到滿足。編寫這本小冊子的目的只是想引發老師們在教授單元一時，在知識上和教學上有更多的討論與探究。

2. 二項展式

一些學生或會覺得二項式定理不只乏味而且難明。乏味的其中一個原因是不了解其用途，除了$(1+x)^n$在迫近的用處外，其他用途尚包括古代中國解方程的想法（即透過完全平方、完全立方等）。就算不談完全平方、立方等構思，略談楊輝（賈憲）三角的故事[1]（從楊輝三角說起）亦應能加添一點趣味。此外，由學生熟知的$(a+b)^2=a^2+2ab+b^2$出發，計算：$(a+b)^3=a^3+3a^2b+3ab^2+b^3$，以至$(a+b)^4$及$(a+b)^5$，從而列出楊輝三角，讓學生觀察當$n$為正整數時，$(a+b)^n$與$(a+b)^{n+1}$展式中，各項係數是如何關聯着的，即$C_r^n+C_{r-1}^n=C_r^{n+1}$，是一個好的切入點。



楊輝三角

至於證明，依法操作並不困難，不過一般學生感到數學歸納法證明二項式定理不只「又長又累贅」，且在日後其他課

1　華羅庚（1964）。《從楊輝三角談起》。北京：人民教育出版社。
梁宗巨（1992）。《數學歷史典故》（第 16 章：賈憲三角—頁 392－407）。沈陽：遼寧教育 出版社。

題均沒用處（考試又通常不會要求考生證明）。其實二項式定理的證明除了數學歸納法外，還可以有較直觀的利用組合想法去證明。就是若

$$(a+b)^n = \underbrace{(a+b)}_{\text{第一個}}\underbrace{(a+b)}_{\text{第二個}} \cdots \underbrace{(a+b)}_{\text{第}n\text{個}}$$

要「生出」$a^r b^{n-r}$來（學生首先要認識到若 a 的冪是 r，b的冪必須是 $n-r$）。相等於從 n 個不同的括號（以次序區分）中選取 r 個來提供 a，其餘的 $n-r$ 個提供 b，故此其方法總數是 C_r^n。這樣，便可推導出二項展式：

$$(a+b)^n = C_0^n a^n b^0 + C_1^n a^{n\text{-}1} b + C_2^n a^{n\text{-}2} b^2 + \cdots + C_r^n a^{n\text{-}r} b^r + \cdots + C_n^n a^0 b^n \qquad (*)$$

因數學歸納法本身也可以進一步淨化數學歸納法思維。若學生覺得 $P(n) \Rightarrow P(n+1)$ 難明，可先證明 $P(3) \Rightarrow P(4)$ 給他們看看，

$$\begin{array}{ll}
 & (a+b)^3 = C_0^3 a^3 + C_1^3 a^2 b + C_2^3 a b^2 + C_3^3 b^3 \\
\times & (a+b) \\ \hline
= & C_0^3 a^4 + C_1^3 a^3 b + C_2^3 a^2 b^2 + C_3^3 a b^3 \\
+ & \qquad\qquad C_0^3 a^3 b + C_1^3 a^2 b^2 + C_2^3 a b^3 + C_3^3 b^4 \\ \hline
= & C_0^3 a^4 + C_1^4 a^3 b + C_2^4 a^2 b^2 + C_3^4 a b^3 + C_3^3 b^4 \\
 & \qquad\qquad\qquad\text{（由於} C_r^n + C_{r-1}^n = C_r^{n+1}\text{）} \\
= & C_0^4 a^4 + C_1^4 a^3 b + C_2^4 a^2 b^2 + C_3^4 a b^3 + C_4^4 b^4 \\
 & \qquad\qquad\text{（由於} C_0^3 = C_0^4 = 1,\quad C_3^3 = C_4^4 = 1\text{）}
\end{array}$$

二項式的展開還有一個有趣的應用，就是求擲數枚骰子擲出指定點數總和的概率。例如擲 3 枚骰子，求點數總和為 10 的形成方法，用類似原理，總數等於$(x + x^2 + \ldots\ldots + x^6)^3$ 中 x^{10} 之係數，上式即

$$\frac{x^3(1-x^6)^3}{(1-x)^3}=x^3(1-3x^6+...)(1+3x+...+36x^7+...)$$

x^{10} 之係數為 $(1)(36)+(-3)(3)=27$。

此外，在表達 (*)時，引入求和記號：

$$(a+b)^n=\sum_{r=0}^{n}C_r^n a^{n-r}b^r,$$

既可讓表達式變得清晰簡潔，亦可令學生掌握應用求和記號的方法與技巧，可謂一舉兩得。

相關網站：

1. 有趣問題

   http://163.21.42.19/mathpath/%B2%C4%A4K%AF%B8/8.htm

2. 二項式與π

   http://www.math.sinica.edu.tw/math_media/d232/23211.pdf

3. 多項式

   http://homepage.ntu.edu.tw/~p94922001/Downloads/Polynomials Coefficient.pdf

4. 楊輝三角

   (a) http://www.math.sinica.edu.tw/math_media/d271/27110.pdf

   (b) http://kss.hkcampus.net/~kss-wsf/theory.htm#triangle

   (c) http://www.chiculture.net/0803/html/c65/0803c65.html

3. 指數及對數函數

對以正實數a為底定義的指數及對數函數，在學生認識指數及對數定律時，便已開始滲入學生的學習中。學生從指數及對數函數的圖像，可以直觀、概略地理解函數的各種性質。然而，細心一想，學生已學習的，其實只是x為有理數時a^x與$\log_a x$的值，而繪製這些函數的圖像，則是通過所謂插值(interpolation)的方法，將指數及對數函數的有理點值，以連續平滑的曲線連接而成的。學生對x為無理數時a^x與$\log_a x$的值，卻沒有認識。這種繪製函數圖像的方法，隱藏着我們的一種信念：指數及對數函數是「連續、光滑」的。事實上，a^x與$\log_a x$在x為無理數時的值，可以利用有理數r逼近x，而以a^r與$\log_a r$的相應逼近值來定義。可以驗證，這樣定義的指數及對數函數是「連續、光滑（可微）」的（詳見附錄三）。至於要詳細有系統的介紹a^x與$\log_a x$，先討論數e，再引入e^x與其逆函數$\ln x = \log_e x$，從而再推廣到一般的a^x與$\log_a x$的做法，不論就歷史發展或實用的角度而言，均屬可取的途徑。

4. e 的簡述

<u>e、指數函數（exponential function）和自然對數（natural logarithm）</u>

e（Euler number）對中學生而言，是一個既神秘，又奇妙的實數。它之所以神秘源於它的定義；它之所以奇妙，來自以它為底（base）所定義的指數函數，以及它相應的逆函數（inverse function）自然對數在物理世界中的種種應用。

<u>e 的引入</u>

很多時可以透過複利的計算公式來引入 e：

$$S = P\left(1+\frac{r}{n}\right)^{nt}$$

此處 P 是本金，S 是本利和，n 是期數，而 r 及 t 則分別代表利率和時期。眾所周知，當 n 增加時，S 將會變大。有趣的問題是：當 n 無窮地遞增時，S 將會如何變化？為使討論更清晰明確，取 $r=t=P=1$，則問題簡化為討論極限

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}$$

是否存在。答案是肯定的，它的值就是 $e = 2.718281828\cdots$。要證明存在極限，我們可確立：

（i） 序列（sequence）$T_n = \left(1+\frac{1}{n}\right)^{n}$ 是嚴格單調上升（strictly monotonic increasing）的，即 $\left(1+\frac{1}{n}\right)^{n} < \left(1+\frac{1}{n+1}\right)^{n+1}$，其中 n 為自然數；

（ii） T_n是有界的，事實上，$2 \le T_n$是明顯的。容易由幾何級數的估值證明對於所有的自然數 n，$T_n \le 3$成立。應用實數序列的單調收斂定理（monotone convergence theorem）：單調上升/下降的實數序列收斂$\Leftrightarrow$該序列為有界序列，從而得知極限$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$存在，故$e$為一有精確意義的數值記號。此外，還可以考慮序列$S_n = 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots\frac{1}{n!}$，$n$ 為自然數，並證明：$e = \lim_{n\to\infty} T_n = \lim_{n\to\infty} S_n$ [2]。

<u>e 的一些重要性質</u>

（1） $2 < e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n < 3$

（2） e是無理數（irrational number），即不能將e寫成$\frac{p}{q}$的形式，其中$p\,,q$為整數，$q \ne 0$

（3） e是超越數（transcendental number），即e不是任何一個以有理數為係數的多項式的根

應用序列S_n與e的誤差估計，可為（2）給出一個漂亮的

[2] 雖然證明是標準的，卻並不容易，直接比較可知$T_n \le S_n$，難的一步是利用二項定理於T_n，n遠大於k ，並看出：

$$\begin{aligned}
T_n &= 1+\sum_{k=1}^{n} C_k^n \frac{1}{n^k} \\
&= 1+\frac{1}{1!}+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\cdots+\frac{1}{k!}\left(1-\frac{1}{n}\right)\times\cdots\times\left(1-\frac{k-1}{n}\right)+\cdots+\frac{1}{n!}\left(1-\frac{1}{n}\right)\times\cdots\times\left(1-\frac{n-1}{n}\right) \\
&\ge 1+\frac{1}{1!}+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\cdots+\frac{1}{k!}\left(1-\frac{1}{n}\right)\times\cdots\times\left(1-\frac{k-1}{n}\right),
\end{aligned}$$

再取$n\to\infty$，推出$e \ge 1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{k!} = S_k$。歸納上述結果，得$T_k \le S_k \le e$，再於上式，應用夾逼定理（squeezing theorem），使$k\to\infty$ 即可得預期結果。

證明[3]，（3）的證明較艱深！不屬於一般的中學課題。

指數函數 e^x 的定義及一些性質

要講解以e為底又或是以一般正實數為底的指數函數，是一個頗令高中老師頭痛的問題。要講解e^x，首先要定義e^x，難題立即湧現：

（I） 當r為有理數時，e^r的意義是熟知的，但當 x 是無理數時，怎樣賦予e^x一個合理的定義？

想不通往往要翻書求救，誰不知在哪些「高等數學」書中，定義以 e 為底的指數函數$e^x:\mathbb{R}\to(0,\infty)$時，往往會參照引入$e$的方法，對實數$x$定義$e^x$為：

（II） $e^x=\lim\limits_{n\to\infty}\left(1+\dfrac{x}{n}\right)^n$，或

（III） $e^x=\lim\limits_{n\to\infty}\sum\limits_{k=0}^{n}\dfrac{x^k}{k!}=\sum\limits_{k=0}^{\infty}\dfrac{x^k}{k!}$

用這些定義來講解，恐怕有點難度。例如，要理解$e^{\sqrt{2}}$時，學生如何去算$e^{\sqrt{2}}=\lim\limits_{n\to\infty}\left(1+\dfrac{\sqrt{2}}{n}\right)^n$或 $\sum\limits_{k=0}^{\infty}\dfrac{\left(\sqrt{2}\right)^k}{k!}$？另外，這樣的定義，會否與一般考慮$e^r$（$r$為有理數）時所用的定義一致？

（IV） 即使老師定義了e^x，接下來要問的，將是如何說服初學者下列有關實數的指數定律，仍然成立：

（1） $e^x\cdot e^y=e^{x+y}$ 及

[3] 參閱復旦大學《數學分析》編寫組（1967）。《數學分析・上冊》。香港：商務印書局。（頁 68－69）。

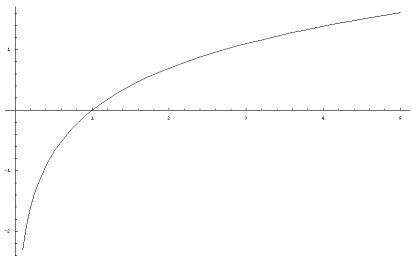
（2） $e^{-x}=\dfrac{1}{e^{x}}$，其中 x 與 y 為實數。

在附錄三中，我們介紹一種方法，在定義 e 之後以 e^{r}（其中 r 為有理數）去逼近 e^{x}，從而使 e^{x} 有意義，並由此看出 e^{x} 是 x 的連續函數以及（1）和（2）成立。

<u>自然對數 ln x 的定義及一些性質</u>

定義了以 e 為底，單調連續的指數函數 $y(x)=e^{x}:\mathbb{R}\to(0,\infty)$ 後，由於它是雙射的，下一步，便可以考慮它的逆函數自然對數函數： $x=e^{y}\Leftrightarrow y=\ln x\ \ (=\log_{e}x)$。

與 $y=e^{x}$ 一樣， $y=\ln x:(0,\infty)\to\mathbb{R}$ 也是單調上升的連續函數。



$y=\ln x$ 的圖像

可驗證，對正實數 x， y 及實數 α，以下對數定律成立：

（1） $\ln x+\ln y=\ln(xy)$，

（2） $\ln x-\ln y=\ln\left(\dfrac{x}{y}\right)$， （3） $\alpha\ln x=\ln x^{\alpha}$ 。

至此，對一般的指數函數 $y(x)=a^{x}$ ， $x\in\mathbb{R}$ ， $a>0$ 均可以通過 $a^{x}=e^{x\ln a}$ 來定義，顯然它是連續函數，且滿足一般的指數定律。

回應指數函數 e^{x} 的定義中的問題（II），要看出 $e^{x}=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{n}$ ， $x\neq0$ ，我們注意到：

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{n}=\lim_{n\to\infty}\left(1+\frac{1}{\frac{n}{x}}\right)^{n}=\lim_{n\to\infty}\left[\left(1+\frac{1}{\frac{n}{x}}\right)^{\frac{n}{x}}\right]^{x}$$

$$=\lim_{n\to\infty}e^{x\ln\left[\left(1+\frac{1}{\frac{n}{x}}\right)^{\frac{n}{x}}\right]}=e^{x\ln\left[\lim_{n\to\infty}\left(1+\frac{1}{\frac{n}{x}}\right)^{\frac{n}{x}}\right]}=e^{x},$$

此處，我們已應用了 e^{x} 與 $\ln x$ 的連續性，以及將在附錄三中證明：

$$（※）\qquad e=\lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^{x}.$$

要想確切理解（III），並不容易，它牽涉到一般的無窮級數理論，在此不作論述。但引導學生從 $e=e^{1}=\lim_{n\to+\infty}\left(1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots\frac{1}{n!}\right)$ 出發去討論 $e^{x}=\lim_{n\to\infty}\sum_{k=0}^{n}\frac{x^{k}}{k!}=\lim_{n\to\infty}S_{n}(x)$ 的意義，卻是一個極具挑戰性的探究活動。對一給定的實數 x，要高中學生證明，（i） $\{S_{n}(x)\}_{n=1}^{\infty}$ 是有界序列，並且(ii)對足夠大的正整數 N，$\{S_{n}(x)\}_{n\geq N}^{\infty}$ 是單調上升序列，從而 $\lim_{n\to\infty}S_{n}(x)$ 存在，是一件非常有意義事！

指數及對數函數的微分

特別重要的是$y=\ln x$的微分性質。對$x>0$應用對數定律，$y=\ln x$ 的連續性及 e 的定義得：

$$(\ln x)'$$

$$=\lim_{h\to 0}\frac{\ln(x+h)-\ln(x)}{h} \qquad \text{（導數的定義）}$$

$$=\lim_{h\to 0}\frac{1}{h}\ln\left(\frac{x+h}{x}\right) \qquad \text{（對數的性質（2））}$$

$$=\lim_{h\to 0}\frac{1}{h}\ln\left(1+\frac{h}{x}\right)$$

$$=\frac{1}{x}\lim_{h\to 0}\ln\left(1+\frac{h}{x}\right)^{\frac{x}{h}} \qquad \text{（對數的性質（3））}$$

$$=\frac{1}{x}\ln\left[\lim_{h\to 0}\left(1+\frac{1}{\frac{x}{h}}\right)^{\frac{x}{h}}\right] \qquad \text{（對數的連續性）}$$

$$=\frac{1}{x}\ln e=\frac{1}{x}. \qquad \text{（由(※)，}\frac{x}{h}\to\pm\infty\text{）}$$

由逆函數的性質$x=\ln(e^x)$，對所有實數x成立，應用鏈式法則（chain rule）得：

$$1=\frac{dx}{dx}=\frac{d[\ln(e^x)]}{dx}=\frac{1}{(e^x)}\frac{d(e^x)}{dx}\ ,$$

指數函數e^x的微分公式：對所有實數x，

$$(*) \qquad \frac{d(e^x)}{dx}=e^x.$$

以一般實數 $a>0$ 為底定義的指數函數 a^x 及對數函數的微分便可推得：

$$(a^x)'=(e^{\ln a^x})'=(e^{x\ln a})'=(\ln a)a^x，$$

$$\text{及}\quad (\log_a x)'=(\frac{\log_e x}{\log_e a})'=\frac{1}{\ln a}(\ln x)'=\frac{1}{x\ln a}。$$

以上是一個有關指數及對數函數的微分公式推導的可行途徑。然而，要推導(*)，只要能證明：

$$(**)\qquad \lim_{x\to 0}\frac{e^x-1}{x}=1$$

便可。細心分析，(**)等價於(*)於 $x=0$ 時成立：

$$\frac{d(e^x)}{dx}\Big|_{x=0}=\lim_{x\to 0}\frac{e^x-1}{x}=1=e^x\Big|_{x=0}。$$

事實上，要確定(**)⇒(*)，只需注意到：

$$\frac{d(e^x)}{dx}=\lim_{h\to 0}\frac{e^{(x+h)}-e^x}{h}=e^x\lim_{h\to 0}\frac{e^h-1}{h}=e^x\cdot 1=e^x.$$

此處可看出(**)作為一個基本的極限關係式，它的功用，便有如 $\lim_{x\to 0}\frac{\sin x}{x}=1$ 在推導三角函數的微分公式時所起的作用一樣，有異曲同工之妙！但要直接由定義出發來導出(**)卻絕非易事。鼓勵學生從 $y=e^x$ 的圖像，動手估算在 $x=0$ 處切線的斜率，又或是直觀地考慮 $e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots$，可知 $\frac{e^x-1}{x}=1+\frac{x}{2!}+\frac{x^2}{3!}+\ldots$ 故當 $x\to 0$，$\frac{e^x-1}{x}\to 1$ 是一個「期望」中的結果。對他們理解(**)是非常有幫助的。

還須指出，以 e 為底的指數函數 $y=e^x$ 還可被定義為函數 $y=y(x)$，它是滿足微分方程：$y'=y$，且適合初始條件 $y(0)=1$

的唯一可微函數 $y = y(x)$。然而，這種引入亦不易為學生所理解。

最後，指數及對數函數除了以上的引入方法外，另一可行的途徑，是在討論了微積分基本定理（fundamental theorem of calculus）後，再引入自然對數函數 $y = \ln x : (0,\infty) \to \mathbb{R}$：

$$\ln x = \int_1^x \frac{1}{t}\,dt \text{ ， } x > 0 \text{ ，}$$

並以其逆函數來定義指數函數 $y = e^x : \mathbb{R} \to (0,\infty)$，其優點是不用以極限來定義冪函數，且容易導出微分及種種特性，詳情可參閱附錄四。

總的來說，引入指數和對數函數，附錄三介紹的方法雖然有點繁複，但能緊扣學生的經驗，易於為學生所接受。附錄四所討論的方法則比較抽象，雖然手法很乾淨利落，但微積分基本定理是關鍵的一步，學生並不容易掌握當中真義！至於利用無窮級數作引入，則最為困難，主要是中學生對無窮級數的認識不足，故不建議採用。當然，教師還有很多「變招」與「方法」。這裏並沒有提及，建議採用與否，最重要的是看教師能否讓學生理解當中的數學內涵，並自圓其說。

對於指數函數 $y = e^x$ 及 $y = \ln x$ 在複利計算、人口增長及放射衰變中的應用，均是熟知的標準課題，在此不贅。

值得注意的是，在應用函數模型：$y = kx^n$ 或 $y = ka^x$ 於實際情況時，往往需要以實驗的方法，從搜集得來有關的 x 及 y 數據，決定 k 及 n(前者)，或 k 及 a(後者)的數值。這都可以通過指數的變換進行。前者，在兩方取自然對數（設：$y, x, k > 0$，否則，可應用如 $(-y) = (-k)x^n$ 的方式再取對數），得 $\ln y = \ln(kx^n) = n\ln x + \ln k$，這是 $(\ln x,\ \ln y)$ 平面上以 n 為斜率，

$\ln k$ 為 y 軸截距的直線方程，由 (x,y) 可知 $(\ln x, \ln y)$，再由數據決定 n 及 $\ln k$ 的數值，亦即 n 及 k 的數值。同理（設：$a,k>0$，$a\neq 1$），在後者兩方取自然對數，得 $\ln y = \ln(ka^x) = (\ln a)x + \ln k$，這是 $(x,\ln y)$ 平面上以 $\ln a$ 為斜率，$\ln k$ 為 y 軸截距的直線方程，由 (x,y) 知 $(x,\ln y)$，再由數據決定 $\ln a$ 及 $\ln k$ 的數值，亦即 a 及 k 的數值。因為我們對尋找通過數據的「最佳直線」方程，所知甚多，故以上的方法，對實際的應用，甚具價值。學生熟習這種確定數學模型中參數的技巧，對日後的發展，十分重要。

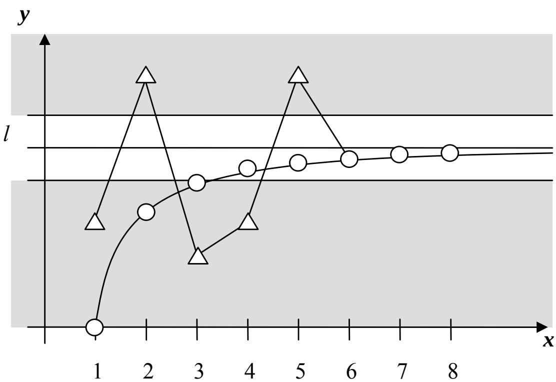
相關網站：

e 的簡述

http://episte.math.ntu.edu.tw/entries/en_e/index.html

5. 極限

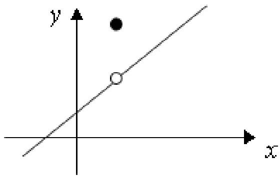
極限的引入通常有三個方向，$\lim_{x\to\pm\infty} f(x)$、$\lim_{x\to a} f(x)$和$\lim_{n\to\infty} a_n$。由於學生大部份曾接觸的函數均是連續，故此除非引用很「怪誕」的函數（如$y = x\sin\frac{1}{x}$之類），否則$\lim_{x\to a} f(x)$均是$f(a)$。故此用$\lim_{x\to a} f(x)$引入極限比較突兀。先以$\lim_{x\to\infty} f(x)$引入極限的概念。再轉到$\lim_{x\to a} f(x)$，然後再引進連續的定義可能比較自然。我們可以利用「最終趨勢」（eventually tendency）的觀念指出若$\lim_{x\to\infty} f(x) = l$，當$x$足夠大時，$f(x)$會「縮進」$l$的左近（隣域），而這些不受前面的數值影響（若以$a_n$而言，那怕你改動前 1000 個$a_n$，也不會影響其極限）。



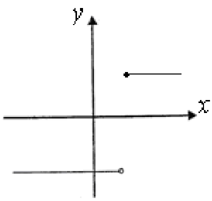
○：$a_n = 1 - \frac{1}{2^n}$

△：改動前面數值不會影響極限

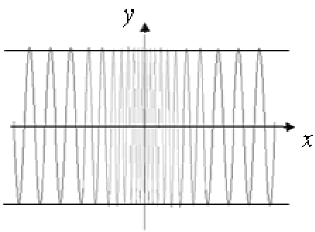
至於$|x|$、sgn(x)等之引入，自然是要利用非連續函數的「非常例子」建構連續的概念，雖然在數量上，非連續函數比連續函數多得多，但對於學生而言，曾經接觸到的非連續函數可以說是絕無僅有。故此引入$|x|$、sgn(x)等可能會有人工化之嫌。我們也許轉一個彎，先看看如果是非連續會出現甚麼情況，可以是「有洞」（所謂「可取代的不連續」－replaceable discontinuity：$\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x)$ 但不等於 $f(a)$）；「落差」（jump：$\lim_{x\to a^+} f(x)$、$\lim_{x\to a^-} f(x)$均存在但 $\lim_{x\to a^+} f(x) \neq \lim_{x\to a^-} f(x)$）或不斷振盪（所謂「實質不連續」－ essential discontinuity：$\lim_{x\to a^+} f(x)$ 或 $\lim_{x\to a^-} f(x)$不存在）。當然我們不必介紹這些名稱的分類法，但從這個入手點再引進$|x|$、sgn(x)等可能更合理一點。

(i) 

$$f(x)=\begin{cases} x+1 & 當 x\neq 1 \\ 4 & 當 x=1 \end{cases}$$

(ii) 

$$f(x)=\begin{cases} 2 & 當 x\geq 1 \\ -2 & 當 x<1 \end{cases}$$

(iii) 

$$f(x)=\begin{cases} 0 & 當 x=0 \\ \sin\frac{1}{x} & 當 x\neq 0 \end{cases}$$

其中一個老師的疑惑是我們不介紹極限的嚴格定義怎樣

可以推導出各種性質？馮振業與黃毅英[4]便提出，先以直觀意念（配以圖像）得出一些基本性質的想法如下：

設 $k,l,m\in\mathbb{R}$，且 $\lim\limits_{x\to a}f(x)=l$ 、 $\lim\limits_{x\to a}g(x)=m$ 均存在，其中 $a\in\mathbb{R}$ 或 $a=\pm\infty$ 則

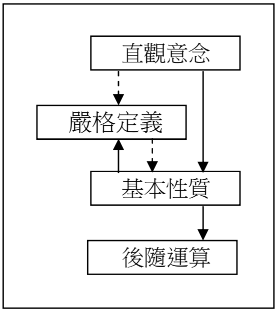
（a） $\lim\limits_{x\to a}k=k$ ；

（b） $\lim\limits_{x\to a}kf(x)$ 存在，且為 kl ；

（c） $\lim\limits_{x\to a}[f(x)\pm g(x)]$ 存在，且為 $l\pm m$ ；

（d） $\lim\limits_{x\to a}f(x)\cdot g(x)$ 存在，且為 lm ；

（e） $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}$ 存在，且為 $\dfrac{l}{m}$，其中 $m\neq 0$ ；



（f） 若 $l=m$ 並在 a 鄰近有 $f(x)\le h(x)\le g(x)$ ，則 $\lim\limits_{x\to a}h(x)=l$ （「三文治原理」）；

（g） 對於正有理數 a ， $\lim\limits_{x\to\infty}\dfrac{1}{x^{a}}=0$ 。

之後便可按照這些基本性質，作出其他推導的運算。

相關網站：

極限簡介

(a) http://www.edp.ust.hk/previous/math/history/5/5_6/5_6_3.htm

(b) http://math.ntut.edu.tw/file/upload/chap1.pdf

(c) http://webcai.math.fcu.edu.tw/calculus/calculus_html/2-2/Limit.htm

(d) http://ind.ntou.edu.tw/~metex/ch2.pdf

4 馮振業、黃毅英（1997）。極限的故事。《課程論壇》7 期，102－105。

6. 求導法

求導法於技術層面並不困難，因為微分比積分「幸運」。透過加法、積、商等法則，不少複雜函數的導數均可「拆開來」求導。再透過鏈式法則，可以「拆括號」。

教師還需注意學生經常犯的「小毛病」，例如在應用「積法則」求導時，學生會否粗心大意的忘記了把乘積中的某一個函數微分；又或在應用「商法則」時，把分子和分母的微分次序弄錯了。提醒學生，以免他們「因小失大」是重要的！此外，有些鏈式法則的記號是十分容易令學生混淆的，例如把鏈式法則寫成：

$\frac{d}{dx}f(g(x))=f'(g(x))g'(x)$。等式右方在 $g'(x)$ 的「 ' 」是表示對變量 x 作微分，而 $f'(g(x))$ 中的「 ' 」則表示對 $f(y)$ 中的變量 y 先作微分，然後再以 $y=g(x)$ 代入所得的結果中。因此，應用繁瑣一點的寫法 $\frac{d}{dx}f(g(x))=f'(y)|_{y=g(x)}\ g'(x)$ 還是較為合適的。另一方面，上式正是「拆括號」的精義所在：左式表示要「進」括號內取 x 的微分，方法正如右式所言，先對 $f(y)$ 中的變量 y 先作微分，再以 $y=g(x)$ 代入；同時微分便進入了括號中，作用於 $g(x)$，所得的數式，就像「鎖鏈」般與 $f'(y)|_{y=g(x)}$ 相扣（乘）着！

順帶一提，雖然學生學過二項展式，要應用它從定義出發去證明導數的非負整數冪法則：$\frac{d}{dx}x^n=nx^{n-1}$ 原非難事，但卻不若應用積法則與數學歸納法（讓學生應用這重要的數學

事實，絕對能使他們得益）的證明來得簡單直接。

相關網站：

求導法簡介

http://www.hkame.org.hk/html/modules/tinyd2/content/Edumath/v16/10LeungCK_diff.pdf

7. 求導法的應用

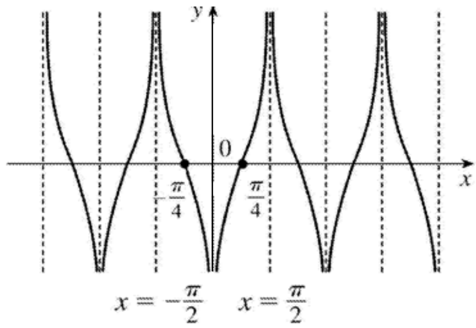
若能夠了解求導法之定義及其歷史根源，其應用也可以不說自明的。不過現時較多集中於求切線，至於變率（rate of change）等較少討論。

至於描繪曲線（包括極大值和極小值的計算），其實有許多討論空間，不必很快的化成機械化的操作。

首先現時電腦軟件這麼方便，何以要費這麼多的功夫去描繪曲線呢？這要牽涉到描繪與逐點繪畫（plot）的分別。逐點繪畫雖然很準確，卻很多時候不如曲線描繪，能描畫出整體圖像的特點。

曲線描繪亦牽涉到微積分與解析幾何的關係。傳統上，高中數學會包括代數、三角、解析幾何和微積分等。現時單元一不包括解析幾何。當然其中一個原因是解析幾何在必修部分已經出現。不過其實亦有另一個觀點認為微積分的學習可以取代了解析幾何的學習。特別是圓錐曲線的探究（甚至有人認為圓錐曲線的學習只不過是基於一些歷史原因，因為古希臘人欣賞圓錐曲線之美）。無論如何，曲線描繪能讓學生跳到一個通則的層面，利用微積分看一般曲線的特性。故此除了逐點去考察曲線的對稱性，截距等，其實可讓學生先作一般性的探討，以增強其觀察能力及對函數的了解，例如

$y = \ln(\tan^2 x)$



一望而知 $f(x) = \ln(\tan^2 x)$ 與 $\tan x$ 一樣是周期函數，並有周期 π 於 $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$。它是偶函數，因為 $f(-x) = \ln(\tan^2(-x))$ $= \ln(\tan^2 x) = f(x)$，故圖像 $y = f(x)$ 以 y 軸為對稱軸，且有垂直漸近線 $x = \pm\frac{\pi}{2}$ 和 $x = 0$。雖然某些特徵不在課程注釋所列出之數點內，但我們卻不會說課程不夠完備。相反，我們要提出的是我們不應墨守成規，在處理問題時不一定要樣板化的逐條核對！

此外，猶如過往純粹數學科的考試報告中指出，考生常常犯極大值就是「$f'(x) = 0,\quad f''(x) < 0$」等毛病。其實極大／小值不一定可導。以往考卷中經常出現。$y = |x|$ 等便是一些關於極大／小值非可導的例子。

例如 1986 年高級程度考試純粹數學科卷 2 第 6 題 $f(x) = x^3 - 3x^2 + 4$，其拐點並不滿足 $f'(x) = 0$。

又有些學生誤以為 $f''(x) = 0$就一定是拐點，但 $f(x) = x^4$有 $f'(0) = f''(0) = 0$，然而 0 並非拐點。

一般而言，若 $f^{(r)}(x)=0\ (1<r<k)$ 而 $f^{(k)}(x)\neq 0$，而若 k 是奇數，則 x 是一拐點。不過這只是充份條件而非必要條件，一函數的拐點甚至可以不能二次可導。1987 年高級程度考試純粹數學科卷 2 第 7 題的 $f(x)=\dfrac{x|x|(x+7)}{x-1}$ 就是了。

以定義而言，拐點是在一個順滑（smooth，即可導）函數中之一點，其切線不只接觸其曲線，且跨越（cross）之[5]。

相關網站：

1. 例子

   http://www.edb.gov.hk/FileManager/TC/Content_4687/differentiation.pdf

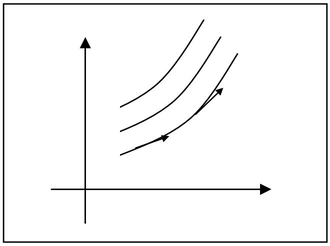
2. 錐術

   http://www.chiculture.net/0803/html/c71/0803c71.html

[5] Wong, N. Y. (1989). Points of inflection. *Mathematics Bulletin,* 18, 19.

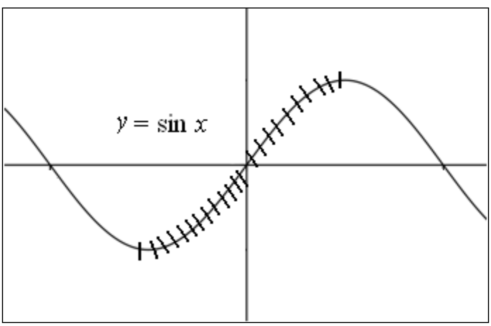
8. 不定積分法

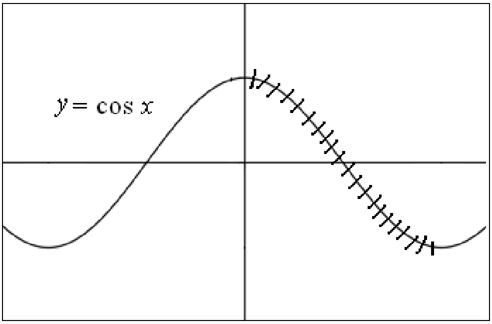
不定積分法之理念也相對簡單。在介紹它是求導法的逆之同時，也可指出當一個函數知道它的趨勢（導數）時，就可「復原」函數本身。但當然這有無限個可能，不過這些可能均只有常數差。在圖像來說，是垂直的平移。

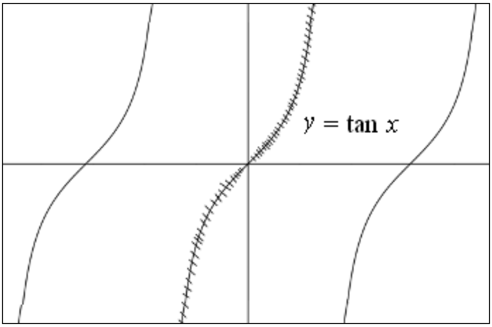


如上所述，積分的運算沒有如導數般有求和、積、商的方便法則，故此較難，一般只能逐類逐類的計算。學習上亦如是，亦可提醒學生某一些積分至今亦未能計出。

至於 $\sin^{-1} x$和 $\tan^{-1} x$的主值，一些學生只靠死記，其實道理很簡單，均是找一個包括原點 0 值的一個連續的一段就是了。







相關網站：

不定積簡介

(a) http://episte.math.ntu.edu.tw/entries/en_integral/index.html

(b) http://episte.math.ntu.edu.tw/articles/mm/mm_02_2_07/index.html

9. 定積分

雖然課程要求介紹微積分基本定理，學生可能有以下疑惑：

（一） 為何突然那麼有興趣研究小片面積的和？

（二） 何以黎曼和（定積分）與導數的逆（不定積分）有關？

這正正是由於（二），我們才可動用所有不定積分的技巧計算定積分，即

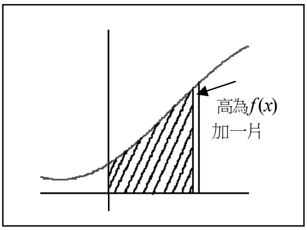
（i） 當 $f(x)$ 為黎曼可積時而 $F(x)$ 是 $f(x)$ 的原函數，即 $F'(x)=f(x)$，有 $\int_a^b f(x)dx = F(x)\Big|_a^b$ 而不必計算複雜的求無窮級數的和：

$$\lim_{n\to\infty}\sum_{r=0}^{n} f(x_r)\Delta x_r$$

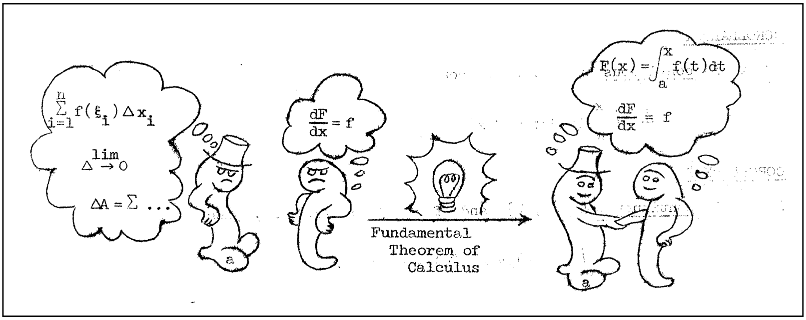
當然微積分基本定理不容易作出證明（在歷史上，I. Barrow：1630 – 1677 雖然有此揣測，但也要等到其學生 I. Newton 才能證出）此外，即使不知道原函數 $F(x)$是否存在，我們仍可證明 $G(x)=\int_a^x f(t)dt$ 是 x 的連續函數，即

$$\lim_{x\to x_o} G(x) = \lim_{x\to x_o}\int_a^x f(t)dt = \int_a^{x_0} f(t)dt = G(x_0)\text{。}$$

（ii）若 $f(x)$是一連續函數時則黎曼和與 $f(x)$有更深一層的關係。以圖像來說，$y=f(x)$ 底下的面積的增加率可以看成再加上一片 $f(x)$。故



此 $\frac{d}{dx}\int_a^x f(t)dt = f(x)$，亦即 $G(x)=\int_a^x f(t)dt$ 是 $f(x)$ 的原函數。



（取自蕭文強教授 1978 年香港大學數學系一年班「基礎數學」講義，謹此致謝）

在討論關於 $\int_a^b f(x)$ 的性質時，可以提醒學生不一定 $b \ge a$。對於 $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$ 也沒有說過 $a \le c \le b$。這就是其中微妙之處，這些性質有不少圖像表示（包括奇函數與偶函數）。但與此同時，學生又應會得跳出圖像的直觀以代數（符號）方式考慮問題。

定積分故然有不少應用，因為不涉及求表面面積，求面積與體積的公式均與直觀吻合，故不應太困難。

函數的對稱性，對定積分的計算，十分重要，例如，當 f 為連續函數，a 為一固定實數時，以下的運算技巧是經常用到的：

(a) 若 f 為奇函數，即對所有 $x \in \mathbb{R}$，有 $f(-x) = -f(x)$，則 $\int_{-a}^{a} f(x)dx = 0$；

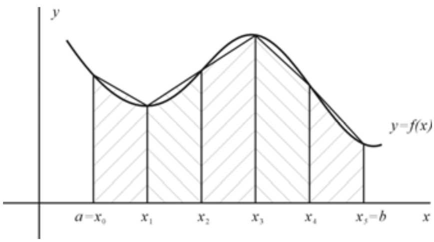
(b) 若 f 為偶函數，即對所有 $x \in \mathbb{R}$，有 $f(-x) = f(x)$，則 $\int_{-a}^{a} f(x)dx = 2\int_0^a f(x)dx$；

(c) 若 f 是以 ω 為週期的週期函數（periodic function），即對所有 $x \in \mathbb{R}$，有 $f(x+\omega)=f(x)$，則 $\int_a^{a+\omega} f(x)dx = \int_0^{\omega} f(x)dx$。

以上事實，可利用適當的變量代換，加以證明。很多時候，繪畫函數的圖像，對幫助學生理解如何找出合適變量代換，是非常有幫助的！

10. 使用梯形法則計算定積分的近似值

用矩形面積的和來逼近一段曲線下方所圍面積，這個基本的想法是定積分中最重要的概念之一，這與在微分法中，用一系列割線來逼近切線以定義函數微分的想法，同樣重要。然而，除了應用矩形外，如下圖所示，如果我們把 $f(x)$的每一小段曲線，換成斜率各不相同的直線，這些斜線顯然比矩形的平頂更近似原來的曲線。



梯形法示意圖

事實上，在圖中梯形的面積可這樣的表示：各梯形左端的高 $L=f(x_{i-1})$ 、右端高 $R=f(x_i)$ 、而寬為 $W=\Delta x_i=x_i-x_{i-1}$ $=\dfrac{b-a}{n}$，故面積為 $(\dfrac{L+R}{2})W=\dfrac{b-a}{n}\left[\dfrac{f(x_{i-1})+f(x_i)}{2}\right]$，考慮圖中各梯形的面積和便可得定積分的梯形近似公式：

$$\int_a^b f(x)dx\approx\frac{b-a}{n}\left[\frac{f(x_0)}{2}+f(x_1)+\cdots+f(x_{n-1})+\frac{f(x_n)}{2}\right]。$$

在應用上述近似公式時，要小心對首尾分點值 $f(x_0)$跟 $f(x_n)$的處理：其餘各分點的值 $f(x_i)$，$1\le i\le n-1$都「整個」的被包括進去了，而 $f(x_0)$跟 $f(x_n)$，則只包含一半而已！

11. 條件概率和獨立性

在討論條件概率和獨立性前，也許我們有需要搞清楚概率的定義。在必修課程中其實也有概率這課題，但礙於此課題佔整個必修課程的篇幅不多，又看似「獨立」於其他課題，關聯性不算高，有關的問題和概念相對地簡單，學生的學習困難似乎多發生在理解問題和誤算事件的可能上，故很多時老師只簡單地介紹：引入概率的古典定義和特性後便很少作深入的討論，學生亦往往因此想不通理論概率（priori probability）和實驗概率（empirical probability）的分別，其實背後大有文章。

我們向學生介紹概率時，通常假設在 n 種等可能方式中，事件 A 有 k 種方式發生，則事件發生的概率以 $P(A)=\frac{k}{n}$ 表示，此古典概率的定義亦存在許多問題。第一，古典概率的定義陷於循環定義的邏輯困局，因事件 A 中有 n 種等可能方式，即表示每個事件發生的概率 $\frac{1}{n}$，利用概率概念定義概率的確存在缺陷。第二，此定義亦沒對 n 附以條件，若 n 為無界時（unbounded），則 $P(A)$的涵意是什麼也不清楚，它只是定性的定義，在數學上並不嚴格，故現代概率論以測度論（measure theory）為基礎，以公設概率（axiomatic probability）代替以往古典學派以相對次數的概念去研究概率，避開了直接討論概率的描述性定義[6]。第三，在現實生活中的實驗往往不能體現

[6] 潘壽山、繆龍驥等（譯）。《數學之內容方法及意義（二）》。台北：財團法人徐氏基金會。

古典概率理論的原因是不能滿足古典概率中要求有限性與等可能性。以擲毫為例，基於毫的兩面並非對稱，不能滿足等可能性，故實際上得公或字的概率不可能是 $\frac{1}{2}$，若想透過此實驗去證明一廂情願的假設是正確更不切實際。

既然古典概率問題很多，為什麼學生還學古典概率的定義呢？原因是此概念簡單、直觀、易明。在假設等可能性的大前提下，計算概率便只需照顧排列與組合的問題。既然古典概率告訴我們擲得公或字的概率應為 $\frac{1}{2}$，但現實經驗卻告訴我們得公或字的概率很多時候不是 $\frac{1}{2}$，究竟兩者能否聯繫？可以的話，又怎樣聯繫呢？事實上擲毫問題已遵守了概率空間的公設（見附錄一）。假設為擲得公的頻數與擲毫次數之比為 $\frac{k}{n}$，p 為擲得公的理論概率的話，根據伯努利（Bernoulli）的弱大數法則（weak law of large number），講出事件 A 的頻率 $\frac{k}{n}$ 依概率收斂（converge in probability）於 p，對任意 $\varepsilon > 0$，即有 $\lim\limits_{n\to\infty} P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = 1$。那麼，當 n 很大（$n \to \infty$）時，$\frac{k}{n}$ 在範圍 $[p - \varepsilon, p + \varepsilon]$ 中的概率為 1。

另一個學生常常討論的詭論（paradoxes）是假設擲一個公正的毫，得公或字的概率是 $\frac{1}{2}$ 的話，那麼若今次擲得公，下次必為字乎？明顯地，答案是否定的。因為每次的擲毫都是獨立但同分佈的隨機過程，第一次擲毫的結果不會影響下次結果，即使擲毫連得 100 次公也不代表第 101 次得字的機會增加。學生可能會問，宏觀地看，連擲 100 次得公或字的機會很小，那又是怎樣的一回事?對於這問題的解釋便得與 Borel

的強大數法則有關。強大數法則是指若 X_i 而 $i=1,\ldots,n$ 是片斷（piecewise）獨立同分佈的隨機變量，同時有 $E(X_i)=p$，當 $n\to\infty$ 則有 $\frac{k}{n}$ 收斂於 p（converges almost surely），即 $P\left(\lim_{n\to\infty}\left|\frac{k}{n}-p\right|=0\right)=1$。強大數法則是討論擲無限多次時，得公的次數與擲毫次數之比是否 $\frac{1}{2}$。對於長遠來說擲得公的次數與擲毫次數之比是否 $\frac{1}{2}$，這個問題視乎樣本空間的條件使得強大數法則成立，詳細的討論可參考有關概率論的文章[7]。

學生在學習完此單元後，很多時只知道計算條件概率的公式是 $P(A\mid B)=\dfrac{P(A\cap B)}{P(B)}$，但利用此公式時卻顯得有點笨拙，甚至根本不知道什麼是條件概率。對於第一個情況的發生很可能是源於學生對理解題意有困難，老師可以做的是盡量解釋題目所涵蓋的資料（無論是表面或隱藏的），至於第二個情況當然也要老師解釋什麼是條件概率。假設 C_1 為樣本空間 S 的子集，我們想得的結果（或我們能知道的結果）是樣本空間 C_1 內的元素。C_2 為樣本空間 S 的另一子集。對於樣本空間 C_1，C_2 內的元素可以定義的話，條件概率就是給定 C_1 時，C_2 在 C_1 假設下的概率。由於 C_2 不一定是 C_1 的子集，所以我們只需考慮 $C_1\cap C_2$ 內的元素，有 $P(C_2\mid C_1)=P(C_1\cap C_2\mid C_1)$。基於此概念，條件概率公式由此產生[8]。

另外，學生在學習條件概率時往往直觀地以為在一般情況下 $P(A\mid B)=P(B\mid A)$，其中 A、B 是兩個不同的事件。若利用

7 蔡聰明（2006）。機率論為何要建立在機率空間上面？《數學傳播》。20 卷 2 期。

8 楊宏章（譯）（1984）。《數理統計》。台北：曉園出版社，頁 67－68。

公式 $P(A \mid B)=\dfrac{P(A \cap B)}{P(B)}$ ，當然可輕易地解釋兩者的分別，但附以具體例子作說明，效果可能更為理想。我們知道當 A 和 B 是獨立事件時，$P(A \mid B)=\dfrac{P(A \cap B)}{P(B)}=\dfrac{P(A)P(B)}{P(B)}=P(A)$。其實學生在處理有關概率的問題時，可多鼓勵他們利用圖像把事件之間的關係形象化，如利用溫氏圖（Venn diagram）去處理互不相容事件，樹形圖（tree diagram）去處理條件概率等。

建議課堂活動：Master Mind

活動目標：學生能明白條件概率的概念。

活動內容：學生 A 需預先設定一行由三個棋子組成的答案，再由學生 B 擺出他估計的答案的排列組合。學生 B 放好棋子後，學生 A 需要就學生 B 的答案給予提示。

相關網站：

1. 有關各類概率的問題與定義

   http://edu.people.com.cn/BIG5/4980701.html

2. 概率與概率法則

   http://www.math.sinica.edu.tw/math_media/episte/episte.phtml?m_id=19106

3. 有趣的例子

   (a) http://zh.wikipedia.org/wiki/%E6%A6%82%E7%8E%87%E8%AE%BA#.E7.94.9F.E6.B4.BB.E4.BE.8B.E5.AD.90

   (b) http://www.math.sinica.edu.tw/math_media/episte/episte.phtml?m_id=23402

12. 貝葉斯定理

由貝葉斯定理所求得之事後概率是由事前概率加上從檢定或樣本中所得之資料結合而成，即可把貝葉斯定理看待成條件概率在一「倒轉」的樹形圖的計算手段，意思是從原先的樹形圖中獲得之樣本資料的事前概率去求得事後概率。貝葉斯定理是找出結果的概率，然後由結果追溯原因，程序與一般概率的計算相反，故可用樣本資料去修正原本概率，對計算有關條件概率的問題時十分有用，因為有時候我們知道一事件結果的概率，而導致此結果的原因有很多可能，而一般概率的加、乘法則有時不足以解決有關的問題，所以貝葉斯定理此時可大派用場。

相關網站：

有關貝葉斯定理的意義

http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=42840

13. 離散隨機變量

隨機變量在實際的教學環境裏所佔的教學時間不多，課程及評估指引中的建議時間是眾多課題中最少，加上這課題幾乎「無數可做」，是否意味着隨機變量相對地不太重要呢？非也。變量是往後所學的根本，為學習概率分佈、期望值、方差和置信區間的前奏，故老師必須向學生解釋清楚變量的概念，否則學生越學越容易迷失，只當變量是計算過程中的元素罷了。

在討論過概率後，為了要深入研究隨機現象，便需要把隨機試驗的結果量化，由討論隨機事件轉移到隨機變量，以便討論和分析其中的規律性，故需引入隨機變量。隨機變量不是自變量，隨機變量的取值是視乎隨機試驗結果而定。一般來說，我們把變量 X 視為函數，X：$S \rightarrow X(S)$，其中 $X(S)$內所有元素是一給定的概率和相加後為 1，隨機變量 X 可分為離散和連續兩類，新高中課程延伸部分單元一內只討論三種離散概率分佈（二項、幾何和泊松分佈）和一種連續概率分佈（正態分佈）。基本上離散和連續變量分別是離散量的樣本空間 S 內元素(樣本點)是有限或可數（countable）的無限，個數值而連續變量的樣本空間內元素（樣本點）是無限個不可數（uncountable）數值。

相關網站：

離散隨機變量概念

http://episte.math.ntu.edu.tw/articles/mm/mm_03_3_07/page5.html

14. 概率分佈、期望值和方差

有時候事件可能發生於一特定頻率，故引入概率分佈（probability distribution）以研究有關現象。概率分佈表示了隨機變量的數值與相應概率的關係。若 $f(x)$為 $X=x$ 的概率時，f 是 X 的概率函數（probability function），其中 $0\le f(x)\le 1$和 $\sum_{x\in S} f(x)=1$ 而 S 為樣本空間。

期望值可以理解成隨機變量 X 的平均數，當然兩者在意義上並不相等。平均數是在得到數據後取所有數值之平均，期望值只是「可能取的數」之平均，是實驗前的合理估計。但把這概念加以擴展，我們很容易理解

$E(X)=\sum_{i=1}^{k} p_i x_i$ （p_i 是數值 x_i的概率）

$=\sum_{i=1}^{k} \frac{f_i}{n} x_i$ （用頻率 $\frac{f_i}{n}$ 取代概率 p_i 其中 $n=\sum_{i=1}^{k} f_i$，當 n 越來越大的，$\frac{f_i}{n}$ 越來越接近 p_i）

$$=\frac{\sum_{i=1}^{k} f_i x_i}{n}$$

$$=\overline{X}$$

其中 $E(X)<\infty$。當然，在碰上 X 是連續隨機變量時，便利用積分去處理，期望值是其中一個了解隨機變量的量。

期望值在初中時已接觸過了，但有很多老師反映較難找到生活上的意義。其實不少決定都運用到期望值的概念，最

為人所熟知的是要不要帶傘子出門。這除了看下雨的概率外，還要參考帶傘子不下雨和沒帶傘子下了雨的「代價」。這些「價值」可以因人而異（例如有些人被雨淋了較易病）、因場合而異（如户外户內），亦要看雨下得多大。

現代風險管理亦按照基本公式「風險=可能性×嚴重性」作出計算。這也可涉及到價值教育，例如說「愛美不愛命」，一些人可以花大量金錢去添置新衣，然而不一定是「美」與「金錢」兩備價值的二分法，例如縱然愛美對於一些人來說未必會用太貴的代價去換取小程度的美服，於是否決這因素又是「美×價錢」這樣的一個乘積。

又例如，如果我們計算賭博的期望值，大部分都是負數的，故長遠而言必成輸家（除非有人認定賭博是一種娛樂，有金錢以外的得益，這又歸到上面「價值」的問題了）。

方差幫助我們了解隨機變數的散佈情況和對集中位置的離散程度，在閱讀一組數據時，除了期望值外，方差越小，數據越集中，反之，數據便越分散。方差的構思是量度數值 X 與平均值之間的差距 $X-E(X)$，差距的平均 $E\left[\left|X-E(X)\right|\right]$ 可表示變數的分散狀況。由於絶對值在運算上的不方便，便以 $E[X-E(X)]^2$ 代替 $E\left[\left|X-E(X)\right|\right]$。在實際的環境裏，利用方差去比較兩組數據時要留意以下兩點：

1. 兩組數據的性質和單位相同，否則比較是沒有意義的。
2. 取值較大的隨機變量之方差通常較大，直接比較兩組取值有明顯差別的數據的方差是不可靠的。故在比較兩組數據前都需要把隨機變量標準化（standardization）。也許學生常常會問為什

麼計算樣本方差的公式是 $\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ 而不是 $\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$。答案是 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是 σ^2 的無偏估量（unbiased estimator）。

$$E(S^2)=E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$$

$$=\frac{1}{n-1}E\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$$

$$=\frac{1}{n-1}E\left(\sum_{i=1}^{n}X_i^{\ 2}-n\bar{X}^2\right)$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i^{\ 2})-nE(\bar{X}^2)\right]$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}[D(X_i)+E(X_i)^2]-n[D(\bar{X})+E(\bar{X})^2]\right\},$$

其中 $D(X)=E\{[X-E(X)]^2\}$

$$=\frac{1}{n-1}\left\{n\sigma^2+n\mu^2-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right\}$$

$$=\frac{1}{n-1}(n-1)\sigma^2$$

$$=\sigma^2$$

但 $E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=E\left(\frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$

$$=\frac{n-1}{n}\sigma^2$$

$$\neq\sigma^2$$

故此，$s^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ 而不是 $\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ 。

建議課堂活動：21 點

活動目標：　學生能明白期望值的概念。

活動內容：　舉出一個具體情境透過概率的計算求出應否再開牌。

相關網站：

1. 期望值
   (a) http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=43423
   (b) http://zh.wikipedia.org/w/index.php?title=%E6%9C%9F%E6%9C%9B%E5%80%BC&variant=zh-tw
2. 方差
   (a) http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=42994
   (b) http://zh.wikipedia.org/w/index.php?title=%E6%96%B9%E5%B7%AE&variant=zh-tw
3. 概率分佈

   http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=43046

15. 二項分佈

二項分佈是由伯努利分佈（或稱二點分佈）延展而得。當重複進行 n 次獨立的伯努利實驗，事件的成功次數 X 便是一個服從於二項分佈的隨機變量。當 $n = 1$，X 便是一個伯努利分佈的隨機變量，故伯努利分佈是二項分佈的特別例子。由於需要考慮事件成功次數的序或排列問題，故在二項分佈的概率密度函數（probability density function）中引入二項係數 C_k^n，其中 n 表示實驗的次數，k 表示事件成功的次數。

如序所言，二項分佈與二項展開式息息相關。若 $n \in \mathbb{Z}^+$，則 $(a+b)^n = \sum_{x=0}^{n} C_x^n a^x b^{n-x}$ 。考慮 $f(x) = \begin{cases} C_x^n p^x (1-p)^{n-x}, & x = 0,1,\ldots,n \\ 0 & \text{，其他} \end{cases}$ 代 $a = p$，$b = 1 - p$，可見二項分佈滿足的條件。

相關網站：

二項分佈

(a) http://episte.math.ntu.edu.tw/articles/sm/sm_16_06_1/index.html

(b) http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=42986

16. 幾何分佈

幾何分佈（Geometric distribution）是負二項分佈的特殊情況。幾何分佈是離散概率分佈。其中一種定義為：在第 n 次伯努利試驗，才得到第一次成功的概率。詳細的說，是：n 次伯努利試驗，前 $n-1$ 次皆失敗，第 n 次才成功的概率。

相關網站：

一些幾何分佈和泊松分佈的有趣例子

http://www.math.sinica.edu.tw/math_media/d252/25211.pdf

17. 泊松分佈

在處理有關二項分佈的問題時，若碰上實驗的次數 n 很大時，計算二項係數和二項概率便變得十分麻煩和複雜，故促使了二項分佈的近似泊松分佈的出現，為二項概率提供近似值。當隨機變量 X 服從於二項分佈 $X \sim \text{Bin}(n，p)$，若 $n \to \infty$，$p \to 0$而 $np \approx \lambda$，其中λ是常數的話，X 便很接近一個泊松隨機變量，因為

$$P(x=k) = C_k^n {p_n}^k (1-p_n)^{n-k}$$

$$= \frac{n(n-1)(n-2)...(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}$$

$$= \frac{{\lambda_n}^k}{k!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdots\frac{n-k+1}{n}\left(1-\frac{\lambda_n}{n}\right)^{n-k}$$

$$= \frac{{\lambda_n}^k}{k!}\cdot\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-k}$$

因為對於確定的 k，

$$\lim_{n\to\infty}{\lambda_n}^k = \lambda^k \text{和} \lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{n-k} = e^{-\lambda}$$

所以 $P(x=k) \to \dfrac{\lambda^k e^{-\lambda}}{k!}$，當 $n \to \infty$。

備註：

1. 記作事件成功的概率，由於與實驗次數有關，故有下標（subscript）n

2. $\lambda_n = n\,p_n$

3. $\lim_{n\to\infty}\lambda_n = \lambda$

泊松分佈除了可作為計算二項分佈的近似值外，有關泊松分佈的問題亦常見於稠密性的問題中，即在特定時間內計算事件出現次數的概率。例如放射性物質的衰變、在三十分鐘內某公司電話來電的次數等等。

泊松分佈是從泊松試驗中得出來，而泊松試驗有以下的特點：

1. 實驗的環境及條件是不變的。
2. 在某一時段或空間內的成功次數是獨立於其他時段或空間的。
3. 在一極短的時段或很細的空間內，成功的概率是與那時段或空間成正比例的，也獨立於那時段或空間以外的成功次數。
4. 在一極短的時段或很細的空間內多於一次成功的概率是很微的，可以算是零。

相關網站：

1. 泊松分佈

   http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=43353

2. 計算機程式

   http://www.mosttss.edu.hk/mathclub/calculator/program.htm

18. 二項、幾何和泊松分佈的應用

教授此課題難度頗高，似乎除了介紹不同應用了各離散的概率分佈例子，亦別無他法，可以教得活潑些是引入有趣和生活化的例子。當學生知道各概率分佈的概念和特質時，可讓他們自行發掘並在課堂中討論。

相關網站：

1. 日常生活中伯努利分佈的例子

   http://zh.wikipedia.org/w/index.php?title=%E4%BC%AF%E5%8A%AA%E5%88%A9%E8%A9%A6%E9%A9%97&variant=zh-hk

2. 二項和泊松分佈之關係的例子

   http://episte.math.ntu.edu.tw/articles/sm/sm_16_07_1/index.html

19. 正態分佈的基本定義及其性質

正態分佈又稱高斯分佈。學習隨機變量分佈是沒有理由錯過正態分佈，因為正態分佈在實踐與理論上同樣重要。除了其應用性高，是日常生活中常遇見的數學模型外，許多離散或連續概率分佈也近似於正態分佈，故在統計學裏，也佔有非常重要的位置。

正態分佈之所以重要是和中心極限定理（central limit theorem）有關。中心極限定理是指一切有關大量隨機變量和極限分佈是正態分佈的定理。直到現在，由中心極限定理已產生出很多獨立同分佈（identical independent distribution）和獨立非同分佈的極限分佈是正態分佈的定理。中心極限定理已成為概率論中極有份量的課題。只要隨機變量為有限值，無論隨機變量的分佈為連續、離散、對稱或偏態，只要樣本數目足夠大的話，中心極限定理均能保證的抽樣分佈（sampling distribution）趨近正態分佈（見附錄二）。

正態分佈的概率密度函數（probability density function）為 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$，$-\infty<x<\infty$。當 $\mu=0$，$\sigma^2=1$時或把其標準化便是標準正態分佈。與其他連續隨機變量一樣，必須滿足 $f(x)\geq 0, \forall x\in\mathbb{R}$ 和 $\int_{-\infty}^{\infty}f(x)dx=1$。第一個非負性條件顯然易見，至於滿足第二個條件的詳情如下：

設 $y=\frac{x-\mu}{\sigma}$

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2}}dy$$

其中，$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\cdot\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{x^2+y^2}{2}}dxdy$

設 $x=r\cos\theta$，$y=r\sin\theta$，可得

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{x^2+y^2}{2}}dxdy=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{0}^{2\pi}re^{-\frac{r^2}{2}}d\theta dr$$

$$=\int_{-\infty}^{\infty}re^{-\frac{r^2}{2}}dr$$

$$=\int_{0}^{\infty}re^{-\frac{r^2}{2}}dr$$

$$=\left[-e^{-\frac{r^2}{2}}\right]_0^{\infty}$$

$$=1$$

$$\therefore\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=1$$

我們可以證明概率密度函數滿足 $\int_{-\infty}^{\infty}f(x)dx=1$。由上面的運算可得知正態分佈的概率密度函數中的常數是 $\frac{1}{\sqrt{2\pi}}$ 的原因是為了使第二個條件得到滿足。由於學生不需要透過積分求概率密度函數因此學生便需透另一渠道去計算 X 的概率了。

若 $X\sim N(\mu,\sigma^2)$，

$$P(a<X<b)=P\left(\frac{a-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{b-\mu}{\sigma}\right)$$

$$=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$$

其中，$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2}}dy$。

當知道 a, b, μ, σ 時，學生便可查表計算其概率。

正態分佈 $N(\mu, \sigma^2)$ 還有以下的特性：

1. 其平均值亦為眾數（mode），即當 $x=\mu$ 時，$f(\mu)$是極大值。
2. 對稱於 $x=\mu$。
3. 這曲線的拐點（inflection point）是在 $x=\mu\pm\sigma$。
4. x 軸是漸近線。

有一點想特別強調，在計算概率分配時，連續隨機變量有少許差異是對於連續隨機變量的定點概率是 0。若 X 是連續隨機變量，則 $P(X<a)=P(X\le a)$換句話，對任意的連續隨機變量是否包含某一點，對計算結果都沒有影響。

相關網站：

1. 正態分佈
   (a) http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=43644
   (b) http://zh.wikipedia.org/wiki/%E6%AD%A3%E6%80%81%E5%88%86%E5%B8%83
2. 中心極限定理
   http://ci58.lib.ntu.edu.tw/cpedia/Content.asp?ID=43655
3. 二項分佈的近似值
   http://www.htjh.tpc.edu.tw/math/LSC/Normal/Normal.htm

20. 抽樣分佈和點估計

抽樣分佈和點估計在舊有的課程已經存在，新課程亦沒有作重大更動，兩者想向學生灌輸以下兩個基本的統計概念。

一、抽樣分佈（sampling distribution）

我們能靠樣本計算統計量，從而推算母體的參數，不同的樣本可能得出不同的統計量，故此可得此統計量的分佈即抽樣分佈。在統計範疇裏，有關由樣本統計量推算母體參數的知識屬於統計推論（statistical inference）。在新課程中，更引入區間估計（總體比例的置信區間和總體平均值的置信區間）對未知參數所在範圍進行估計。講授抽樣分佈時便需向學生介紹概率分佈、隨機抽樣、驗抽樣分佈等概念，從中引入中心極限定理，當然能配合具體的實解說對建立概念會更有幫助。

二、無偏估計量（unbiased estimator）

當了解抽樣理論後，便需要利用統計推斷的方法實在地處理如何從眾樣本中推斷有關母體的訊息。在得到樣本估計量後，推斷母體參數前，我們需要檢定估計量的 1. 無偏性（unbiasedness），2. 效率性（efficiency），3. 一致性（consistency）。有一點我們需要學生留意的是估計如 $\bar{x}$, s^2等，它們的數值是會隨着不同的樣本而有所不同估計，所以我們特別有興趣知道估計量的無偏性，即 $E(\hat{\theta})=\theta$。

21. 總體平均值的置信區間

雖然總體平均值的置信區間與舊有數學與統計比較是新引入的課題，但其實只是點估計的延續。點估計是以一個統計量去估計母體的參數（population parameter），可以想像點估計難免有誤差，而這誤差亦難以估計，故利用區間估計去指出參數的可能範圍是比較可行的做法。如想進一步知道參數在特定可能範圍的概率，或估計的誤差，這便與置信區間有關了。設 θ 是總體的參數，如果統計量 $\hat{\theta}_1,\hat{\theta}_2$ 滿足 $P(\hat{\theta}_1<\theta<\hat{\theta}_2)=1-\alpha$，其中 $\alpha\in\mathbb{R}$ 和 $0<\alpha<1$，則隨機區間（$\hat{\theta}_1,\hat{\theta}_2$）是 θ 的 $(1-\alpha)$ 置信區間，區間的長短反映了估計的精確度，α 的大小顯示了估計的可靠性。如果 θ 是母體的平均值，（$\hat{\theta}_1,\hat{\theta}_2$）便是總體平均值的置信區間。順帶一提，對於給定的置信水平 $(1-\alpha)$，置信區間不是唯一的。即滿足 $P(\hat{\theta}_1<\theta<\hat{\theta}_2)=1-\alpha$ 的 $\hat{\theta}_1,\hat{\theta}_2$ 是可以有多於一組的解。由於新課程只考慮母體的數量是無限或有限但抽取是有放回的情況，樣本數目足夠大或知道抽樣分佈近似於正態分佈，故總體平均值的置信區間應為 $(\bar{x}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\bar{x}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$。若總體的標準差 σ 無法得知，我們可以 σ 的估計 s 代替。$\frac{\sigma}{\sqrt{n}}$ 是點估計量的標準誤差，$z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ 是抽樣誤差，$z_{\frac{\alpha}{2}}$ 是臨界值（critical value）。平均值的 $(1-\alpha)$ 置信區間來自

$$1-\alpha=P(-z_{\frac{\alpha}{2}}<\theta<z_{\frac{\alpha}{2}})$$

$$= P(-z_{\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < z_{\frac{\alpha}{2}})$$

$$= P(\mu - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \overline{X} < \mu + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}})$$

$$= P(\overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}})$$

置信區間的範圍主要受兩個因素影響。第一是樣本的大小。當樣本越大，點估計和抽樣的誤差便越小，所得之置信區間越狹窄。第二是置信水平（confidence level）。為了建立更高的置信度即減低誤差，讓參數在置信區間內的機會提高，故置信區間範圍便需擴大。若想精確度和可靠性兩者兼得、取得平衡，那就必需加大樣本數目。

正如前文所說，這置信區間並非唯一，只要我們取不對稱的雙α 點時，同樣地$P\left(-z_{\frac{4\alpha}{5}} < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < z_{\frac{\alpha}{5}}\right) = 1 - \alpha$ 成立。由此，亦可得另一置信區間。雖然理論上在同一置信水平可得出不同的置信區間，但實際上仍會選取以對稱的雙α 點為置信區間，因為其所求得的置信區間長度比其他以不對稱求得的置信區間長度為短，其精確度較高，所以我們通常由對稱的雙α 點求置信區間。

有一點必須提醒學生的是μ 的（$1-\alpha$）100%置信區間不是表示μ 有（$1-\alpha$）100%的概率發生在此範圍內，而是所有（$1-\alpha$）100%置信區間中有（$1-\alpha$）100%的置信區間會包含μ。由於μ 的置信區間(μ_1, μ_2)是根據 P（$\mu_1 < \mu < \mu_2$）$=1-\alpha$ 計算出來，即不同的樣本會計算出不同的(μ_1, μ_2)，故此(μ_1, μ_2)亦

是隨機變量，我們只能說 (μ_1, μ_2) 能被正確地估計的概率為（1－α）100%，但不能說 μ 在 (μ_1, μ_2) 的概率為（1－α）100%，因為 μ 已有一個固定值（只是未知），故 μ 是否在 (μ_1, μ_2) 內並非一個隨機變量。

22. 總體比例的置信區間

總體比例的置信區間的學理大部分與總體平均值的置信區間相似，兩者之分別是前者估計的是母體內特定事件的成功比例，參數 p 是比例或概率，其中 $0 < p < 1$。即若 X 是某事件的成功次數，則 X 服從於二項分佈（$X \sim \text{Bin}(n, p)$）。由於估計總體比例的統計量 p 服從於二項分佈，故其標準差則為 $\sqrt{\dfrac{p(1-p)}{n}}$。我們知道當樣本數目足夠大時，二項分佈亦近似於正態分佈，相似地，總體比例的置信區間為 $\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\right)$。在計算總體比例的置信區間時，我們以 $\hat{p} = \dfrac{X}{n}$ 作為總體比例的點估計。由於 $\dfrac{X - np}{\sqrt{np(1-p)}}$ 和 $\dfrac{X - np}{\sqrt{n\left(\dfrac{X}{n}\right)\left(1 - \dfrac{X}{n}\right)}}$ 之極限分佈[9]（limiting distribution）為 $N(0, 1)$，且 $\hat{p}$ 依概率收斂於 p；1－$\hat{p}$ 依概率收斂於 1－p。若對任意 $\varepsilon > 0$，即有 $\lim_{n\to\infty} P\left(\left|\dfrac{X}{n} - p\right| < \varepsilon\right) = 1$。所以有概率近似值

$$1 - \alpha = P\left(-z_{\frac{\alpha}{2}} < \frac{X - np}{\sqrt{n\left(\frac{X}{n}\right)\left(1 - \frac{X}{n}\right)}} < z_{\frac{\alpha}{2}}\right)$$

9　楊宏章（譯）（1984）。《數理統計》。台北：曉園出版社，頁 226－227。

$$= P\left(-z_{\frac{\alpha}{2}} < \frac{\frac{X}{n} - p}{\sqrt{\frac{1}{n}\left(\frac{X}{n}\right)\left(1 - \frac{X}{n}\right)}} < z_{\frac{\alpha}{2}} \right)$$

$$= P\left(-z_{\frac{\alpha}{2}} < \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < z_{\frac{\alpha}{2}} \right)$$

$$= P\left(p - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < \hat{p} < p + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

$$= P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

參考書目

[1] Casella, G, & Berger R. L. (2002). *Statistical inference*. Australia: Thomson Learning.

這部書適合老師作為在概率與統計知識上有更深入的鞏固與增潤之讀物。

[2] Dretzke, B. J., & Heilman, K. A. (1998). *Statistics with Microsoft Excel*. Upper Saddle River, New Jersey: Prentice Hall.

此書提供大量利用試算表（Excel®）找各分佈的臨界值和概率的例子，亦有教授利用試算表（Excel®）計算置信區間。若老師想以資訊科技輔助教授統計和概率，書中內容對老師很有幫助。

[3] Freund, J. E., & Perles, B. M. (2007). *Modern elementary statistics* (12^{th} edition). Upper Saddle River, New Jersey: Pearson Prentice Hall.

此書通常為大學用書，但值得向學生推薦。書中的內容淺白易明，提供大量範例和練習題，是作為補充教材的一個好選擇。

[4] Gnedenko, B. V. (1997). *Theory of probability* (6^{th} edition). Amsterdam: Gordon and Breach.

[5] 李佑義（譯）(1986)。概率論淺說。北京：宇航出版社。中譯本的篇幅不長，大部分章節與課程有關，對概率的加法與乘法定理有仔細的解釋。另外，第十一章（英文

版為第七章）的大數定理亦不容錯過。

[6] Jagdish, K. P., & Campbell, B. R. (1996). *Handbook of the normal distribution* (2nd edition). New York: Marcel Dekker.

看見書名，大概可以想象內容主要討論正態分佈。介紹此書是希望老師特別留意第六章和第七章的中心極限定理。書中以不同方法解釋利用正態分佈是二項分佈的近似值。雖然解說談不上非常詳盡，但很少書會同一時間以幾個策略作解釋，作為開闊眼界和加深對有關課題的認識，是一個不錯的選擇。老師也要有一個心理準備，其數學內容不是短時間內可以讀得明白。

[7] Ross, S. (2006). *A first course in probability* (7th edition). Upper Saddle River, New Jersey: Pearson Prentice Hall.

一本不可不看的作品，基本上大部分在概率學上應懂的知識均已涵蓋。雖然是大學用書，但內容淺白易明。特別推薦第八章中中心極限定理之證明，是眾多版本中比較簡潔利落的一個。

[8] Scheaffer, L. R. (1996). *Activity-based statistics: Instructor resources*. New York: Springer-Verlag.

香港坊間對教授統計和概率的教材套不多，特別是高中。此書的教學活動，雖然不是全部活動都與課程有關，但當中有很多是合適的，例如幾何分配、置信區間、中心極限定理等。書中有 48 個不同的活動，每個活動均有詳細說明活動的目標、所需教具、學生的先備知識、活動程序、部分更有工作紙和引導問題，絕對是老師在教學上不可缺少的參考書目。

[9] Shaughnessy, J. M. (2005). *Statistical questions from the classroom*. Reston, Virginia: National Council of Teachers of Mathematics.

全書有十一個章節，每章節均解答一個統計問題，而這些問題正是學生在學習統計的過程中常遇到的問題。大部分的問題都是直接與課程有關，例如： what is a p value? Why do we divide by $n-1$ instead of n? 等。由於整章節以解答一個問題為主，再輔以例子補充，故當中解說十分詳盡，是老師在教學上不可錯過的讀物。

[10] Smith, G. (1998). *Introduction to statistical reasoning*. Boston, Massachusetts: WCB/McGraw-Hill.

亦是另一本常用的大學用書，內容涉及統計學的層面較廣，但其賣點是書中穿插很多貼近生活的日常例子，雖然例子中的事件非取材自香港，但一樣容易產生共鳴。

[11] The School Mathematics Project. (1991). *Living with uncertainty*. Cambridge, U.K.: Cambridge University Press.

這本書的對象是 16－19 歲的讀者，雖然對香港的高中學生來說相對淺易，但所引用的例子卻相當好玩有趣，部分例子便可變成課堂活動或遊戲，值得老師花時間一看。

[12] 蕭文強、林建（1982）。《概率萬花筒》。香港：廣角鏡出版社。

個人一向喜歡蕭教授的著作，原因是選材有趣和生活化，數學味濃，無論讀者的數學程度高或低，總覺得好看。此書亦言，同時已把不同的概率問題分類，對老師利用書中例子在教學上作舉隅時更見方便！

[13] 李華剛、廖俊傑、邵慰慈（1999）。《統計學入門》。香港：香港教育圖書公司。

[14] 蘇國樑（2003）。《統計學（第 2 版）》。台北縣蘆洲市：國立空中大學。

[15] 郭信霖、許淑卿（1999）。《統計學》。台北：華立圖書股份有限公司。

三本書的風格相似，易讀易明，對統計和概率的基本概念以精簡的篇幅把重點扼要地描述。由於全書只討論重點，節奏明快，讀起來有暢快之感，故實在適合學生閱讀。

[16] 鄭惟厚（2007）。《你不能不懂的統計常識》。台北：天下文化。

鄭教授以有趣的問題作引入，解讀在應用層面常面對的統計問題。若老師在課堂上想舉些日常生活例子引起同學的學習動機，此書是一個好選擇。

[17] 孫榮恆（2004）。《趣味隨機問題》。北京：科學出版社。

全書有百多個概率問題，如書名一樣，全都是好玩的數學。當中不乏中外的經典數學名題。可是並非所有問題都在課程內，老師需要在其中細意挑選。

[18] 汪振鵬（1999）。《概率與數理統計（第 2 版）》。上海：華東師範大學出版社。

[19] 何燦芝（1984）。《概率統計學習指導》。長沙：湖南科學技術出版社。

這兩部書對概率的概念和定義有比較準確的解釋與描述。

[20] 陳明哲（1972）。《排列組合》。台北：中央書局。

[21] 陳明哲（1980）。《機率》。台北：中央書局。

[22] Li, K. C. (1983). *Probability* (3rd edition). Hong Kong: Educational Publishing House.

附錄一：弱大數定律

要明白弱大數定律之前，首先要明白何謂「概率上收斂」（converge in probability）。

若 X_n 為隨機變量之序列。X 又是一隨機變量。

我們稱 $X_n \xrightarrow{P} X$（概率上收斂）若

$$\lim_{n\to\infty} P(\left|\overline{X_n} - X\right| \geq \varepsilon) = 0 \text{，} \quad \forall \varepsilon > 0 \text{。}$$

弱大數定律是這樣說的：

若 X 是一隨機變量，$X_1, X_2, \ldots,\ X_n$ 為 X 之隨機抽樣，那末 $\overline{X_n} \xrightarrow{P} \mu$，其中 μ 為 X 之平均值。

故此，這定律並沒有說 $\overline{X_n} \longrightarrow \mu$，只是說 $\overline{X_n} \xrightarrow{P} \mu$。換言之，不是說當 n 愈大，樣本平均值 $\overline{X_n}$ 會愈來愈近 μ。再換句話說，就是這定律並不是說，當樣本愈來愈大（sample size），樣本平均值會趨向 μ。

我們可以這樣理解。首先以擲毫為例（假設得正反兩面的概率均等）。擲毫這事件形成的隨機變量稱為 X。我們作出抽樣 $X_1, X_2, \ldots,\ X_n$。假如得了

0 1 1 0 0 0 1 0 1 0 0 1 1 1 1 0 ……

當 n 愈來愈大時，得出了不同的 $\overline{X_n}$：

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	…	…
X_n	0	1	1	0	0	0	1	0	1	0	0	1	1	1	1	0	…	…
$\overline{X_n}$	0	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{4}$	$\frac{2}{5}$	$\frac{2}{6}$	$\frac{3}{7}$	$\frac{3}{8}$	$\frac{4}{9}$	$\frac{4}{10}$	$\frac{4}{11}$	$\frac{5}{12}$	$\frac{6}{13}$	$\frac{7}{14}$	$\frac{8}{15}$	$\frac{8}{16}$	…	…

在這特別的情況，X_n確是愈來愈近$\frac{1}{2}$。但現並不表示這是必然的（這就是與常見的理解之最大分別）。

我們確有可能出現「特異」的情況，（硬幣正反兩面的概率同樣均等如）如

$$1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\cdots\cdots$$

或

$$1\,1\,1\,1\,1\,1\,0\,1\,1\,1\,1\,1\,1\,1\,1\,1\cdots\cdots$$

只是這種出現的可能性比較低（這就是所謂「概率上收斂」）。

我們可以這麼理解（以下已是「大數定律」之註釋，而非大數定律本身：因為大數定律並沒有如以下例子中「找不同的人做實驗」的含意）。假若我們叫很多學生（例如說 100 個學生）擲同一個硬幣（亦是假設了得正反兩面概率均等）許多許多次（例如說 1000 次）。每個學生就得出了一個數值。當擲的次數愈多（1000 次增多，而不是 100 增多）。這 100 個數值會有相當多是接近$\frac{1}{2}$的（但仍有可能有部分偏離$\frac{1}{2}$的）。

附錄二：中心極限定理

前文提及，中心極限定理是指一切有關大量隨機變量和極限分佈是正態分佈的定理，由於只想介紹部分常見的中心極限定理，着重提出它的應用與意義，故不加以證明。

中心極限定理

設 $X_1, X_2, \ldots\ldots, X_n$ 為獨立同分佈的隨機變量（各有期望值 $\mu < \infty$ 及方差 $\sigma^2 > 0$）和 $Y_n = \dfrac{(X_1 + \ldots + X_n) - n\mu}{\sigma\sqrt{n}}$，則 $\lim\limits_{n\to\infty} P(Y_n < x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ 即 $\dfrac{(X_1 + \ldots + X_n) - n\mu}{\sigma\sqrt{n}}$ 有極限分佈標準正態分佈。

一、獨立同分佈的中心極限定理

定理一： 設 $X_1, X_2, \ldots\ldots, X_n$ 為互相獨立的隨機變量序列，各有期望值 μ 及方差 σ^2，當 $n \to \infty$ 時，則 $Y_n = \dfrac{(X_1 + \ldots + X_n) - n\mu}{\sigma\sqrt{n}}$ 的分佈趨於標準正態分佈，即 $\lim\limits_{n\to\infty} P(Y_n < x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ 此定理為計算獨立同分佈隨機變量之和的近似概率提供方便之門。

定理二： （棣莫弗 — 拉著拉斯極限定理）設 μ_n 是重覆 n 次伯努利實驗中事件 A 發生的次數，而 p 是事件 A 在每次實驗中發生的概率，對任意 x，則 $\lim\limits_{n\to\infty} P\left(\dfrac{\mu_n - np}{\sqrt{npq}} < x\right) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$，其中 $q = 1 - p$ 由於

$\mu_n \sim \text{Bin}(n, p)$，所以定理二說明了二項分佈近似於正態分佈。現實中，只要 n 的數目足夠(n >15)和 p 接近 $\frac{1}{2}$ 時效果已相當顯注，若 p 太接近 0 或 1 時，則需增加樣數目才可有良好效果。基於二項分佈是離散分佈和正態分佈為連續分佈對於定點概率的計算之不同，故以正態分佈去計算二項分佈的近似概率的需要作出修正。

建議課堂活動：概率雨實驗（彈珠機）

活動目標：　　利用實驗展示二項分佈近似於正態分佈。

活動內容：　　利用不同數目的彈珠，模擬重覆伯努尼試驗趨近常模分析。

二、獨立非同分佈的中心極限定理

定理三：（李雅普諾夫定理 Lingdeberg）設 $X_1, X_2, \ldots\ldots, X_n$ 是互相獨立的隨機變量，其中 $E(X_i) = \mu_i$，$Var(X_i) = \sigma_i^2$，i=1, 2, … , n，而 $B_n^{\ 2} = \sum_{i=1}^{n} \sigma_i^{\ 2}$，若存在正數 ε，使得 $\lim_{n\to\infty} \frac{1}{B_n^{\ 2+\varepsilon}} \sum_{i=1}^{n} E\,|\,x_i - \mu_i\,|^{2+\varepsilon} = 0$，則對任何實數 x，

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n}(x_i - \mu_i)}{B_n} < x \right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

附錄三：指數函數 e^x 性質的論述

就學生的學習經驗而言，如前所闡釋，學生已知 e 是一正實數，要回應第 8 頁的問題（I），比較自然的想法是從有理數冪的 e^r（r 為有理數）過渡到實數冪的 e^x。這只需要考慮 x 是無理數的情況便已足夠。

由於無理數 x 可以寫成無限非循環的十進制小數，我們可依此選取兩列單調有理數列 $\{x^-{}_n\}$ 及 $\{x^+{}_n\}$ 使得

(i) 對所有的自然數 n，$x^-{}_n < x < x^+{}_n$；

(ii) $\{x^-{}_n\}$ 單調上升，$\{x^+{}_n\}$ 單調下降；

(iii) $\lim_{n\to\infty} x^-{}_n = \lim_{n\to\infty} x^+{}_n = x$。

事實上，由 x 的無限非循環十進制小數表示，我們可用「捨尾法」及「進一法」選取序列：

$x^-{}_n$ 為 x 精確至第 n 位小數的不足近似值；

$x^+{}_n$ 為 x 精確至第 n 位小數的過剩近似值。

例如：$x = \sqrt{2} = 1.414213562\ \cdots$ [10]

$\{x^-{}_n\}$ 為 1.4，1.41，1.414，1.4142， 1.41421，…，↑(單調遞增)

$\{x^+{}_n\}$ 為 1.5，1.42，1.415，1.4143， 1.41422，…。↓(單調遞減)

[10] 又當 $x = -\sqrt{2} = -1.414213562\ \cdots$ 時

$\{x^-{}_n\}$ 為 -1.5， -1.42， -1.415， -1.4143， -1.41422，…，

$\{x^+{}_n\}$ 為 -1.4， -1.41， -1.414， -1.4142， -1.41421，…。

考慮實數數列 $\{e^{x^-{}_n}\}$ 及 $\{e^{x^+{}_n}\}$。由有理序列的單調性，及指數的有理冪性質知：對於所有自然數 n， $e^{x^-{}_1} \le e^{x^-{}_n} < e^{x^+{}_n} \le e^{x^+{}_1}$ 再由實數序列的單調收斂定理得知 $\alpha = \lim\limits_{n\to\infty} e^{x^-{}_n}$ 及 $\beta = \lim\limits_{n\to\infty} e^{x^+{}_n}$ 存在。考慮有理序列 $\{x^-{}_n\}$ 及 $\{x^+{}_n\}$ 的構作，得 $x^+{}_n - x^-{}_n = 10^{-n}$，從而有

$$\beta - \alpha = \lim_{n\to\infty} e^{x^+{}_n} - \lim_{n\to\infty} e^{x^-{}_n} = \lim_{n\to\infty} e^{x^-{}_n}(e^{x^+{}_n - x^-{}_n} - 1)$$
$$= \lim_{n\to\infty} e^{x^-{}_n}(e^{10^{-n}} - 1) = \lim_{n\to\infty} e^{x^-{}_n} \cdot \lim_{n\to\infty}(e^{10^{-n}} - 1) = \alpha \cdot 0 = 0$$ [11]

我們定義：

$$e^x = \lim_{n\to\infty} e^{x^-{}_n} = \lim_{n\to\infty} e^{x^+{}_n}\text{。}$$

由先前的論述，對所有自然數 n，我們有如下不等式成立：

$$e^{x^-{}_n} \le e^x \le e^{x^+{}_n}$$ [12]。

可再進一步，由此驗證 e^x 的定義與選取的逼近序列無關，以上做法應可滿足（I）的需求。

事實上，可依上述方法，定義以實數 x 為冪的 a^x， $a > 0$，從而，我們可使 $\left(1+\dfrac{1}{x}\right)^x$ 有確切的意義，並證明：

(※) $$e = \lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^x\text{。}$$

這裡特別要指出，對任意實數 x，在沒有把 $\left(1+\dfrac{1}{x}\right)^x$ 的意義弄清楚以前，用以下步驟去「論證」$x\to\infty$ 的(※)是有欠穩妥的：

11　此處我們應用了熟知的結果：$\lim\limits_{m\to\infty} a^{\frac{1}{m}} = 1, a > 0$，m 為自然數。

12　這對一般以 $a > 1$ 為底的實數成立，若 $0 < a < 1$，則 $a^{x^-{}_n} \le a^x \le a^{x^+{}_n}$

設 $[x]$ 是唯一的整數使得 $[x] \le x < [x]+1$ 成立。考慮 $x \to \infty$ $\Rightarrow [x] \to \infty$，故可選 $1 \le [x]$，使得

$$1 \le [x] \le x < [x]+1 \Rightarrow \frac{1}{[x]+1} < \frac{1}{x} \le \frac{1}{[x]} \Rightarrow 1+\frac{1}{[x]+1} < 1+\frac{1}{x} \le 1+\frac{1}{[x]}$$

$$\Rightarrow \left(1+\frac{1}{[x]+1}\right)^{[x]} < \left(1+\frac{1}{x}\right)^{x} \le \left(1+\frac{1}{[x]}\right)^{[x]+1}$$

（這其實有欠穩妥者：上式中央的一項未有定義，另外，不等式是否成立仍屬未知！但按照之前的方法定義一般正數的實數冪，則毋須為此擔憂。）

由夾逼定理（Squeezing Theorem），

$$\lim_{x\to+\infty}\left(1+\frac{1}{[x]+1}\right)^{[x]} = \lim_{x\to+\infty}\frac{\left(1+\frac{1}{[x]+1}\right)^{[x]+1}}{\left(1+\frac{1}{[x]+1}\right)} = e \text{ 及}$$

$$\lim_{x\to+\infty}\left(1+\frac{1}{[x]}\right)^{[x]+1} = \lim_{x\to+\infty}\left(1+\frac{1}{[x]}\right)^{[x]}\left(1+\frac{1}{[x]}\right) = e \quad \text{得(※)！}$$

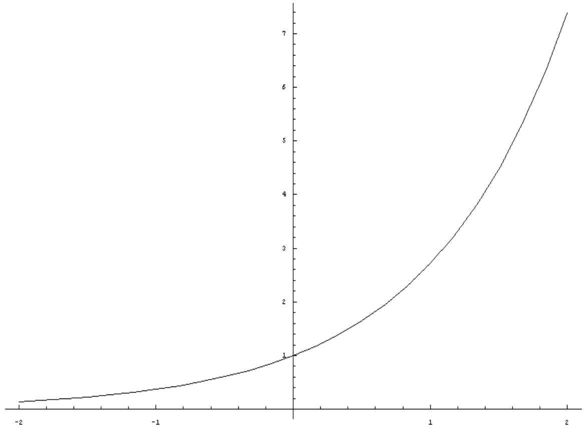
另一方面，還可以考慮 $m=-n$，計算得

$$\lim_{n\to-\infty}\left(1+\frac{1}{n}\right)^{n} = \lim_{m\to\infty}\left(1+\frac{1}{m-1}\right)^{m} = \lim_{m\to\infty}\left(1+\frac{1}{m-1}\right)^{m-1}\cdot\lim_{m\to\infty}\left(1+\frac{1}{m-1}\right) = e$$ 。

同樣，還可以證明：$\lim_{t\to0}(1+t)^{1/t} = e$。

交代過（I）以後，最後，就第 8 頁的問題（IV）而言，應用所提供有關 e^x 的定義，有理冪 e^r 的指數定律及極限的性質，便可順利推導上述的實數指數定律（1）及（2）成立；還可證明 e^x 對 x 是單調遞增且連續的，並有 $\lim_{x\to\infty} e^x = \infty$ 及 $\lim_{x\to-\infty} e^x = 0$，從而得到單調連續的 $y(x)=e^x$ 圖像，以及 $y(x)=e^x : \mathbb{R} \to (0,\infty)$ 是

雙射（bijective）的。在此，我們不再贅述了。



$y(x)=e^{x}$的圖像

附錄四：引入指數與對數函數的另一方法

本附錄將以黎曼積分為基礎，首先以定積分定義自然對數，證明它所滿足的對數定律。之後，通過考慮自然對數的逆函數，便可獲得以e為底的指數函數。我們除了證明指數函數滿足的實數指數定律外，更討論指數與對數函數的微分性質。我們將假定讀者熟知基礎的黎曼積分理論如：連續函數在閉區間上的可積性、定積分對上下限之連續性、以及微積分基本定理（Fundamental Theorem of Calculus）等標準結果。

<u>自然對數的定義及性質</u>

設 $\mathbb{R}^+$為正實數集，函數 $f:\mathbb{R}^+\to\mathbb{R},\ \ f(x)=\dfrac{1}{x},\ \ x\in\mathbb{R}^+$是單調下降且連續的，故此，以下述定積分刻畫之函數有意義。

<u>定義 1</u>　　以定積分定義函數 $\ln:\mathbb{R}^+\to\mathbb{R}$，使得對 $x\in\mathbb{R}^+$

$$\ln\ x\ =\ \int_1^x \frac{1}{t}\,dt\quad 。$$

此函數稱為自然對數，或在內容清晰的情況下，簡稱為對數。在證明下列定理，細緻地確立 $\ln:\mathbb{R}^+\to\mathbb{R}$ 的基本性質後，函數為什麼要以「對數」為名，便將清楚明白了。

<u>定理 2</u>　　以下有關函數 $\ln:\mathbb{R}^+\to\mathbb{R}$ 的命題成立：

（1）　$\ln(1)=0$.

（2）　對任意 $a,b\in\mathbb{R}^+,\ \ \ln(ab)=\ln(a)+\ln(b)$.

（3）　對任意 $a,b\in\mathbb{R}^+,\ \ \ln(\dfrac{a}{b})=\ln(a)-\ln(b)$.

（4） 對任意 $a \in \mathbb{Q}$ 及 $b \in \mathbb{R}^+, \ln(b^a) = a\ln(b)$.

（5） 函數 ln 是嚴格單調遞增及連續的.

（6） 函數 ln 是滿射（surjective or onto）.

（7） 對任意 $x \in \mathbb{R}^+, \frac{d}{dx}\ln(x) = \frac{1}{x}$.

證明

（1）是明顯的，因為 $\ln(1) = \int_1^1 \frac{1}{t}dt = 0$，對任意 $a,b \in \mathbb{R}^+$，設 $g(x) = ax$，則由定積分的變量代換得：

$$\begin{aligned}
\ln(ab) &= \int_1^{ab} \frac{1}{x}dx = \int_1^a \frac{1}{x}dx + \int_a^{ab} \frac{1}{x}dx \\
&= \ln(a) + \int_a^{ab} \frac{1}{ax}adx \\
&= \ln(a) + \int_a^{ab} \frac{1}{g(x)}g'(x)dx \\
&= \ln(a) + \int_1^b \frac{1}{t}dt = \ln(a) + \ln(b)
\end{aligned}$$

故（2）得證。

應用（2），考慮 $\ln(a) = \ln(\frac{a}{b}\cdot b) = \ln(\frac{a}{b}) + \ln(b) \Rightarrow$（3）。

（4）可直接由（1）－（3）推得，只需注意對任意自然數 n，（2）$\Rightarrow \ln(b^n) = n\ln(b)$ 以及 $\ln(b^{\frac{1}{n}}) = \frac{1}{n}\ln(b)$。

要證（5），考慮 $a < b$，其中 $a,b \in \mathbb{R}^+$

$$\ln(b) = \int_1^b \frac{1}{t}dt = \int_1^a \frac{1}{t}dt + \int_a^b \frac{1}{t}dt > \int_1^a \frac{1}{t}dt = \ln(a)$$

故此 ln 是嚴格單調遞增函數，其連續性則是黎曼積分上下限

連續性的直接推論。

要證（6），設 n 為正整數，$n>1$，考慮在 $[1,n]$ 上的函數 $y=\frac{1}{x}$，由函數嚴格單調遞減的性質知：

$$\int_1^n \frac{1}{x}\,dx = \sum_{k=1}^{n-1}\int_k^{k+1}\frac{1}{x}\,dx \geq \sum_{k=1}^{n-1}\frac{1}{k+1}\int_k^{k+1}dx = \sum_{k=2}^{n}\frac{1}{k}$$

由於 $\sum_{k=2}^{\infty}\frac{1}{k}$ 發散於正無窮，故對於任意 $M\in\mathbb{R}^+$，可找到對應的 n 使得 $\ln(n)>M>0=\ln(1)$。應用連續函數的中間值定理（Intermediate Value Theorem）於 $[1,n]$ 上，知存在 $x\in(1,n)$ 使得 $\ln(x)=M$。再由（1）及（3）知 $\ln(1)=0$ 及 $\ln(\frac{1}{x})=-\ln(x)$，從而函數 $\ln:\mathbb{R}^+\to\mathbb{R}$ 為滿射。

最後，因 $y=\frac{1}{x}$ 在 n 為正整數的區間 $[1,n]$ 上連續，故（7）是微積分基本定理的直接結果。定理得證。

由（5）$(\Rightarrow 1-1)$ 及（6）$(\Rightarrow onto)$ 知 $\ln:\mathbb{R}^+\to\mathbb{R}$ 為雙射（bijective）的，故其逆函數存在。

<u>定義 3</u>　定義 $\exp:\mathbb{R}\to\mathbb{R}^+$ 為函數 $\ln:\mathbb{R}^+\to\mathbb{R}$ 的逆函數。稱此函數為指數函數（exponential function）。

以下定理總結了函數 exp 的幾項重要的性質。

<u>定理 4</u>　以下有關函數 $\exp:\mathbb{R}\to\mathbb{R}^+$ 的命題成立：

（1）設 e 為正實數，使得 $1=\ln(e)$，則對所有自然數 n，有 $\exp(n)=e^n$

（2）應用（1）中所定義的正實數 e，對所有自然數 n，有 $\exp\left(\frac{1}{n}\right)=\sqrt[n]{e}$

（3） 對任意 $a,b\in\mathbb{R}^+$， $\exp(a+b)=\exp(a)\exp(b)$

（4） 對任意 $a,b\in\mathbb{R}^+$， $\exp(a-b)=\dfrac{\exp(a)}{\exp(b)}$

（5） $\exp:\mathbb{R}\to\mathbb{R}^+$是單調遞增及連續的雙映射（bijective map）

（6） 對任意 $x\in\mathbb{R}$， $\dfrac{d}{dx}\exp(x)=\exp(x)$

（7） $\lim\limits_{n\to\infty}\left(1+\dfrac{1}{n}\right)^n=e$

證明

函數 $\exp:\mathbb{R}\to\mathbb{R}^+$的性質（1）－（5），可由逆函數ln的對應性質直接推得。

由於 $\ln(\exp(x))=x$對所有 $x\in\mathbb{R}$ 成立，故

$\dfrac{d}{dx}\ln(\exp(x))=1 \quad\Rightarrow\quad \dfrac{1}{\exp(x)}\dfrac{d}{dx}\exp(x)=1$從而（6）得證。

要證（7），由 $\dfrac{d}{dx}\ln(x)=\dfrac{1}{x}$ 知 $\left.\dfrac{d}{dx}\ln(x)\right|_{x=1}=1$

即 $\lim\limits_{h\to0}\dfrac{\ln(1+h)-\ln(1)}{h}=1$，

故得：

$$\lim_{h\to0}\frac{\ln(1+h)-\ln(1)}{h}=1\Rightarrow\lim_{n\to\infty}\frac{\ln\left(1+\frac{1}{n}\right)}{\frac{1}{n}}=1$$

（由極限的性質，當中 $h=\dfrac{1}{n}$）

$$\Rightarrow \lim_{n\to\infty} \ln\left(1+\frac{1}{n}\right)^n = 1$$ （由 ln的性質）

（由 ln的連續性）

$$\Rightarrow \ln\left(\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\right) = 1$$ （由 ln的可逆性）

$$\Rightarrow \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$$

故原定理得證。

由定理 2 的（4）知：對 $a \in \mathbb{Q}$ 及 $b \in \mathbb{R}^+$ $b^a = \exp(a\ln(b))$成立。自然地，我們可沿用這樣的方式去定義以一般的，以正實數為底的指數函數：

<u>定義 5</u>　對 $a \in \mathbb{R}$ 及 $b \in \mathbb{R}^+$，定義 b^a 為

(#) $$b^a = \exp(a\ln(b))$$

這樣，由定義 5，我們看到 $e^x = \exp(x)$對所有 $x \in \mathbb{R}$ 成立。故我們可視函數 exp確是以 e 為底的指數函數！由比，交替應用 e^x 與 $\exp(x)$ 記號亦不會導致含混了。另外，容易驗證 $\ln(b^a) = a\ln(b)$亦成立，$\ln x$與 $\log_e x$亦不用再區分了。

應用(#)，函數 exp及 ln的性質，我們便可確立以任意 $b \in \mathbb{R}^+$為底的實數指數定律：

<u>定理 6</u>

（1） 對任意 $a_1, a_2 \in \mathbb{R}$，$b^{a_1} \times b^{a_2} = b^{a_1+a_2}$

（2） 對任意 $a \in \mathbb{R}$，$\dfrac{1}{b^a} = b^{-a}$

（3） 對任意 $a_1, a_2 \in \mathbb{R}$，$(b^{a_1})^{a_2} = b^{a_1 a_2}$

最後是一般微分公式：

若 $a \in \mathbb{R}$， $x \in \mathbb{R}^{+}$

$$(\dagger) \qquad \frac{d}{dx}x^{\alpha} = \frac{d}{dx}\exp(\alpha\ln(x)) = \exp(\alpha\ln(x)) \times \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

這跟初學微積分時經常用到的公式 $\frac{d}{dx}x^{n} = nx^{n-1}$，其中 n 為正整數及 $x \in \mathbb{R}$，完全一致！

此外，當底數 $k \in \mathbb{R}^{+}$ 為固定值，指數 $x \in \mathbb{R}$ 作為變量時的微分公式亦可輕易求得：

$$(\ddagger) \qquad \frac{d}{dx}k^{x} = \frac{d}{dx}\exp(x\ln(k)) = \exp(x\ln(k)) \times \ln(k) = \ln(k)k^{x}$$

附錄五：離散概率分佈的期望值和方差

雖然課程中並未要求學生就各離散概率分佈的期望值與方差作計算，但當有需要向學生解釋各分佈的期望值與方差時，附錄五希望可以為老師在教學上或在本科知識上有增潤作用。

二項分佈

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{n} xC_x^n p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} xC_x^n p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= n\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= np\sum_{x=1}^{n} C_{x-1}^{n-1} p^{x-1} (1-p)^{n-x} \\
&= np\sum_{y=0}^{n-1} C_y^{n-1} p^y (1-p)^{n-1-y}, \text{ 其中 } y = x-1 \\
&= np
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{n} x^2 C_x^n p^x (1-p)^{n-x} \\
&= np\sum_{x=1}^{n} xC_{x-1}^{n-1} p^{x-1} (1-p)^{n-x} \\
&= np\sum_{y=0}^{n-1} (y+1)C_y^{n-1} p^y (1-p)^{n-1-y}, \text{代 } y = x-1 \\
&= np\sum_{y=0}^{n-1} yC_y^{n-1} p^y (1-p)^{n-1-y} + np\sum_{y=0}^{n-1} C_y^{n-1} p^y (1-p)^{n-1-y} \\
&= np[(n-1)p] + np
\end{aligned}
$$

$$\begin{aligned}\mathrm{Var}(X) &= np[(n-1)p] + np - (np)^2\\ &= n^2p^2 - np^2 + np - n^2p^2\\ &= np\,(1-p)\end{aligned}$$

幾何分佈

$$\begin{aligned}E(X) &= \sum_{x=1}^{\infty} xpq^{x-1}\\ &= p\sum_{x=1}^{\infty}\frac{d(q^x)}{dq}\\ &= p\frac{d}{dq}\left(\sum_{x=1}^{\infty}p^x\right)\\ &= p\frac{d}{dq}\left(\frac{q}{1-q}\right)\\ &= p\cdot\frac{1}{(1-q)^2}\\ &= \frac{1}{p}\end{aligned}$$

$$\begin{aligned}E(X^2) &= \sum_{x=1}^{\infty} x^2pq^{x-1}\\ &= p\sum_{x=1}^{\infty} x(x-1+1)q^{x-1}\\ &= pq\sum_{x=1}^{\infty} x(x-1)q^{x-2} + \sum_{x=1}^{\infty} xpq^{x-1}\\ &= pq\sum_{x=1}^{\infty}\frac{d^2(q^x)}{dq^2} + \frac{1}{p}\\ &= pq\frac{d^2}{dq^2}\sum_{x=1}^{\infty}q^x + \frac{1}{p}\\ &= \frac{2pq}{(1-q)^3} + \frac{1}{p}\\ &= \frac{2-p}{p^2}\end{aligned}$$

$$\mathrm{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

泊松分佈

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!} \\
&= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!}, \text{ 其中 } y = x - 1 \\
&= \lambda e^{-\lambda} e^{\lambda} \\
&= \lambda
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^{x}}{x!} \\
&= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{x\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{(y+1)\lambda^{y}}{y!}, \text{代 } y = x - 1 \\
&= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{y\lambda^{y}}{y!} + \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!} \\
&= \lambda(\lambda + 1)
\end{aligned}
$$

$$
\mathrm{Var}(X) = \lambda(\lambda + 1) - \lambda^2 = \lambda
$$

數學百子櫃系列	作者
（一）漫談數學學與教－新高中數學課程必修部分	張家麟、黃毅英、韓藝詩
（二）漫談數學學與教－新高中數學課程延伸部分單元一	韓藝詩、黃毅英、張家麟
（三）漫談數學學與教－新高中數學課程延伸部分單元二（即將出版）	黃毅英、張家麟、韓藝詩
（四）談天說地話數學	梁子傑
（五）數學的應用：圖像處理－矩陣世紀	陳漢夫
（六）數學的應用：投資組合及市場效率	楊良河
（七）數學的應用：基因及蛋白的分析（即將出版）	徐國榮