Journal of Propagations in Probability and Statistics
Vol. 2, No. 2, February 2002
Far Eastern ed., pp. 141−146

統 計 薪 傳
第二卷　第二期
中華民國 91 年 2 月
遠東版 141−146 頁

微積分與中央極限定理

王思博
義守大學

摘　要

統計學中最著名且最重要的定理，大概就是中央極限定理了。許多大專學生雖然學習過這個定理，但卻不知如何證明此定理。本文針對中央極限定理的證明以及相關的預備知識做一個總整理，以供修習統計學之大專學生課外閱讀以及統計學教師做為輔助教材之用。

關鍵詞：中央極限定理，泰勒定理，動差母函數，分佈收斂。

□民國九十年十月收稿，九十一年一月修訂。
□本文作者為輔仁大學夜間部統計系民國八十九學年畢業生，現為義守大學資訊管理研究所研究生。

1. 前言

記得在大學時期修習微積分課程時，老師曾經說過，微積分有三大定理：微積分基本定理(Fundamental Theorem of Calculus)，平均值定理(Mean Value Theorem)與泰勒定理(Taylor's Theorem)。這三大定理不僅在數學課程中極爲重要，同時亦可被廣泛地應用於機率及統計學的理論與方法，而其中泰勒定理則與統計學裡赫赫有名的中央極限定理(Central Limit Theorem，以下簡稱 CLT)有很直接且重要的關係。凡是修過統計學的大專學生，不論唸什麼科系，大概都或多或少地學習過 CLT。但這些曾經學習過 CLT 的眾多大專學生裡，能夠非常清楚明白且深入地了解該定理者，卻是鳳毛麟角；尤其若是談到該定理之證明，那麼很可能在 100 位隨機抽樣之曾經修習過統計學的大專學生裡，至少有 98 位都是"莫宰羊"呢！說實話，一位不是資質特優且好學不倦的一般大專學生，若沒有受教於認真負責的好老師，而想要徹底學會 CLT 並了解該定理之證明，的確不是容易的事。其原因之一，則是大多數的大專學生在修習微積分課程時，沒有學會甚至不曾學過泰勒定理，那麼在數學基礎知識不足之情況下，想要了解 CLT 之證明，當然是困難重重了。筆者在大學時期很幸運地受教於一位認真負責的好老師，並修習他開授之微積分與數理統計課程。老師曾經針對 CLT 作清楚的說明

及整理，筆者至今依然受用良多，因而撰寫這一篇學習心得報告，希望能將 CLT 的證明以及相關的預備知識做一個總整理，使莘莘學子能真正了解 CLT 該如何證明及其與微積分的關係。

2.中央極限定理之證明

2.1 預備知識

CLT 之證明，牽涉到不少有關機率與微積分的預備知識。首先我們來了解一些在機率方面的預備知識如下：

<u>分佈收斂的定義</u>：假設隨機變數序列Y_n的CDF序列為$G_n(y)$, $n = 1, 2, 3,\ldots$ 。若存在一隨機變數Y，其CDF為$G(y)$，滿足

$$\lim_{n\to\infty} G_n(y) = G(y) \text{ ，}$$

則稱$G(y)$為$G_n(y)$的極限分佈(limiting distribution)，或稱Y_n分佈收斂至Y (Y_n converges in distribution to Y)，記為 $Y_n \xrightarrow{d} Y$ 。

<u>動差母函數之定義及常用基本公式與定理</u>：設X為一連續型隨機變數，其機率密度函數(probability density function，以下簡稱 pdf)為 $f_X(x)$ ，則定義

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx$$

為隨機變數 X 的動差母函數(moment generating function，以下簡稱 MGF)，其中 $M_X(t)$ 括號內的 t 為一"數學變數"，因此 $M_X(t)$ 是一個數學函數，而下標 X 為一隨機變數，表示動差母函數 $M_X(t)$ 是由隨機變數 X 所產生。若 X 為一離散型隨機變數，則將上述定義中之積分符號改為連加(summation)符號如下

$$M_X(t) = E(e^{tX}) = \sum_{\text{all } x} e^{tx} f_X(x)$$

即可。動差母函數之常用基本公式與定理如下：

1) $E(X^k) = M_X^{(k)}(t)\big|_{t=0}$ 其中 $M_X^{(k)}(t)$ 代表 $M_X(t)$ 的 k 階微分。 (2-1)

2) $M_{aX}(t) = M_X(at)$ ，a 為任一實數。 (2-2)

3) 獨立隨機變數之和的動差母函數定理：設 $X_1, X_2, \ldots, X_n$ 為相互獨立之隨機變數且 $Y = \sum_{i=1}^{n} X_i$ ，則 $M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t)$ 。 (2-3)

4) 動差母函數的唯一性(Uniqueness)定理：設 X_1 與 X_2 二隨機變數之CDF分別為$F_1(x)$和$F_2(x)$，而其MGF分別為$M_1(t)$和$M_2(t)$。若 $F_1(x) = F_2(x)$，則對於所有在開區間$(-h, h)$內之t，$h > 0$，等式$M_1(t) = M_2(t)$成立；反之，若已知$M_1(t) = M_2(t)$，則$F_1(x) = F_2(x)$。此定理的意義是：一個隨機變數和它的MGF之間的關係是一對一的，亦即一個隨機變數只有一個MGF(若存在的話)。反之，若已知某MGF，則該MGF僅代表某"唯一"之隨機變數。因此，一個隨機變數和它的MGF之間的關係，就好像一個人和此人的指紋之間的關係，或好像

一個人和其身分證號碼之間的關係一樣。此定理之證明極爲艱深，牽涉到 Laplace transform之理論。有興趣瞭解此定理之證明的讀者，不妨先閱讀 Widder (1961)一書之 13、14 兩章，然後再參閱Chung (1974)、Feller (1971)二書及Curtiss (1942)一文。

5) 動差母函數之收斂定理：設$Y_1, Y_2, \cdots$爲一組隨機變數序列，而與其對應之 CDF 序列及 MGF 序列分別爲$G_1(y), G_2(y), \cdots$及$M_1(t), M_2(t), \cdots$。設隨機變數 Y 之 CDF 及 MGF 分別爲 $G(y)$及 $M(t)$，又對於所有在開區間$(-h, h)$內之 t，$h>0$，等式$\lim_{n\to\infty} M_n(t) = M(t)$成立，則$\lim_{n\to\infty} G_n(y) = G(y)$，其中 y 爲 $G(y)$之任一連續點(continuity point)。此定理的意義是：隨機變數序列之分佈收斂，就相當於其對應的 MGF 序列之收斂。此定理之證明亦十分艱深，讀者不妨參閱 Lukacs (1970)一書及 Curtiss (1942)一文。

接著，我們來了解一些在微積分方面的預備知識如下：

<u>泰勒定理(Taylor's Theorem)</u>：設I爲包含a的一個開區間，而函數$f(x)$在此開區間I內可微分至第$(n+1)$階，亦即$f^{(n+1)}(x)$存在，$\forall x \in I$。 則對於開區間I中之任一$x \neq a$而言，必存在某嚴格介於a與x間之數字t_x，使得

$$\begin{aligned} f(x) = f(a) &+ \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots \\ &+ \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(t_x)}{(n+1)!}(x-a)^{(n+1)} \end{aligned}$$

以上泰勒定理中若令 $n = 0$，則該定理退化爲眾所周知的平均值定理(Mean Value Theorem)。因此，吾人可將泰勒定理視爲一個推廣型態的平均值定理。至於泰勒定理之證明，則在一般的微積分書籍中均可查閱得到，此處不再贅述。而其證明中的一個重要關鍵，則應用到下述定理：

<u>洛氏定理(Rolle's Theorem)</u>：設函數$f(x)$在封閉區間$[a, b]$內爲連續，同時在開區間(a, b)內爲可微分，又$f(a) = f(b)$，則存在某數字$c \in (a,b)$使得$f'(c) = 0$。

以上洛氏定理中若將“$f(a) = f(b)$”之假設條件去除，則定理之結論變爲“存在某數字$c \in (a,b)$使得$f(b) - f(a) = f'(c)(b-a)$，而該定理就變成爲眾所周知的平均值定理。因此，吾人可將洛氏定理視爲一個特殊狀況之下的平均值定理。至於洛氏定理之證明，則在一般的微積分書籍中亦均可查閱得到，此處不再贅述。而其證明中的一個重要關鍵，則應用到下述定理：

<u>極值定理(Extreme Value Theorem或 The Max-Min Theorem)</u>：設函數$f(x)$在封閉區間$[a, b]$內爲連續，則$f(x)$ 在$[a, b]$內具有極大值和極小值，即在$[a, b]$內存在兩個位置c_1與c_2，使得$f(c_1) \le f(x) \le f(c_2)$，$\forall x \in [a,b]$。

極值定理之內容雖然簡單易懂，但其證明可並不簡單，要在數學分析(mathematical analysis)的書籍中較易查閱得到，例如 Belding and Mitchell (1991)一書之 89 頁。最後，介紹一個與極限有關之有用公式如下(請參閱 Bain and Engelhardt 1992 一書 234 頁)：

$$\lim_{n\to\infty}\left[1+\frac{c}{n}+\frac{d(n)}{n}\right]^{nb}=e^{cb}\ \text{ if }\ \lim_{n\to\infty}d(n)=0 。\qquad (2\text{-}4)$$

此公式之證明可從 Rohatgi (1976)一書之 278 頁查閱得到。具備了以上的諸多預備知識之後，要證明中央極限定理，就不是難事了。

2.2 證明中央極限定理

現在我們來看看如何證明 CLT。首先我們要了解，CLT 依其假設條件之不同而有諸多不同的版本。本文所討論之版本，其假設條件爲：母體無限大(infinite population)，母體平均數、變異數、動差母函數皆存在，而樣本爲 i.i.d. (independent and identically distributed)。這個版本是一般數理統計學書籍中最常見的版本，可能也是最容易證明的版本，其內容與證明如下：

<u>中央極限定理</u>：設 $X_1, X_2, \ldots, X_n$ 爲自某無限大母體抽出之一組隨機樣本(即i.i.d. 樣本)，其母體平均數 μ 、變異數 σ^2 、動差母函數 $M_X(t)$ 皆存在(故 $\sigma^2<\infty$)。令

$$Z_n=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}=\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\qquad \text{其中}\quad \overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i ,$$

則 $Z_n \xrightarrow{d} Z \sim N(0,1)$，即 Z_n 分佈趨近於標準常態分佈。

<u>證明</u>：設 $m(t)$ 爲 $X-\mu$ 的MGF，即 $m(t)=M_{X-\mu}(t)$ ，其中 X 爲代表母體分佈之隨機變數。由於 $m(0) = 1$ (Why？)，又根據前一小節之公式(2-1)，可得

$$m'(0)=E(X-\mu)=E(X)-E(\mu)=\mu-\mu=0 \quad \text{及} \quad m''(0)=E(X-\mu)^2=\sigma^2 。$$

利用泰勒定理將 $m(t)$ 以 $n = 1$ 及 $a = 0$ 展開，則存在某嚴格介於 0 與 t 間之數字 ξ (即 $0<|\xi|<|t|$)，使得

$$m(t)=m(0)+m'(0)t+\frac{m''(\xi)t^2}{2}=1+\frac{m''(\xi)t^2}{2}=1+\frac{\sigma^2t^2}{2}+\frac{[m''(\xi)-\sigma^2]t^2}{2}$$

接著，因爲

$$\sum_{i=1}^{n}X_i-n\mu=(X_1+X_2+\cdots+X_n)-(\mu+\mu+\cdots+\mu)=\sum_{i=1}^{n}(X_i-\mu) ,$$

而其中之 $(X_i-\mu)$ (共 n 個)爲 i.i.d.(Why？)，故由前節公式(2-2)與定理(2-3)可得

$$M_{Z_n}(t)=M_{\sum(X_i-\mu)}\left(\frac{t}{\sqrt{n}\sigma}\right)=\prod_{i=1}^{n}M_{X_i-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)=\left[M_{X-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n$$

$$=\left[m\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n=\left[1+\frac{\sigma^2}{2}\left(\frac{t}{\sqrt{n}\sigma}\right)^2+\frac{[m''(\xi_1)-\sigma^2]}{2}\left(\frac{t}{\sqrt{n}\sigma}\right)^2\right]^n$$

$$=\left[1+\frac{t^2}{2n}+\frac{[m''(\xi_1)-\sigma^2]t^2}{2n\sigma^2}\right]^n$$

其中 $0<|\xi_1|<|t|/\sqrt{n}\sigma$。請留意此處之 ξ_1 並不是之前 $m(t)$ 的泰勒展式中之 ξ，因為此處之動差母函數 $m(\cdot)$ 其括號中的數學變數已經改變為 $t/\sqrt{n}\sigma$。令

$$d(n)=\frac{[m''(\xi_1)-\sigma^2]t^2}{2\sigma^2}\text{，}$$

則

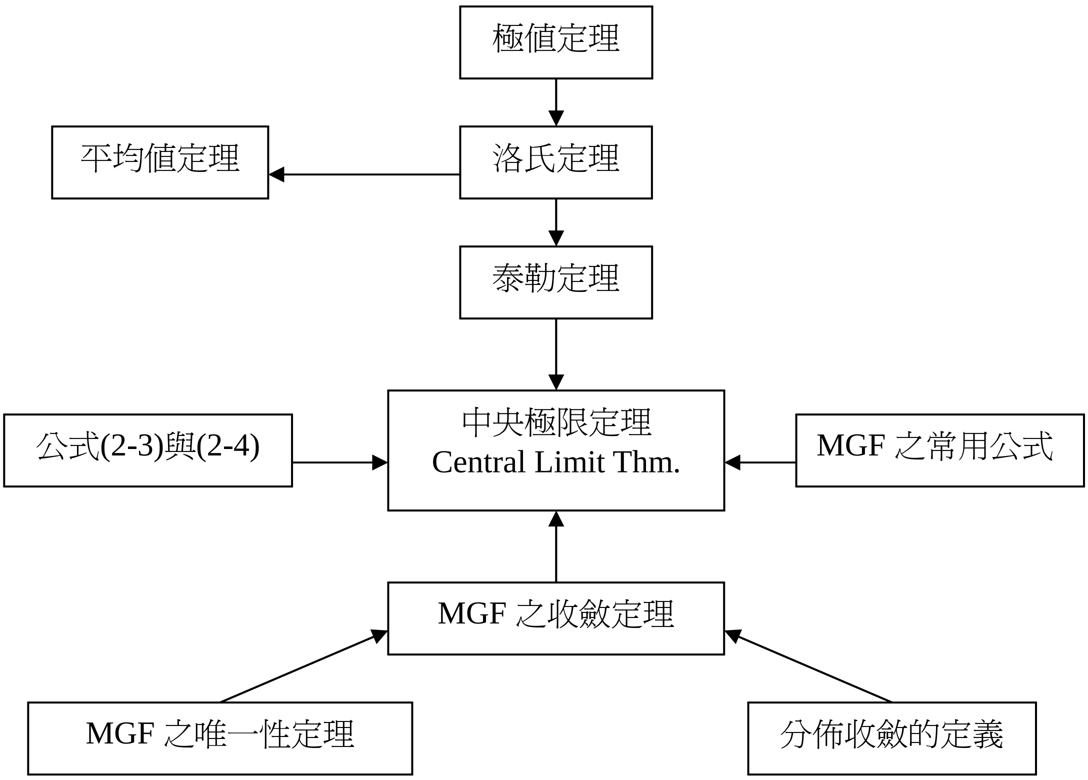
"$n\to\infty \Rightarrow (t/\sqrt{n}\sigma)\to 0 \Rightarrow \xi_1\to 0 \Rightarrow m''(\xi_1)\to m''(0)=\sigma^2 \Rightarrow m''(\xi_1)-\sigma^2\to 0$"，

故 $\lim_{n\to\infty} d(n)=0$。因此，由前一小節之公式(2-4)得知

$$\lim_{n\to\infty} M_{Z_n}(t)=\lim_{n\to\infty}\left[1+\frac{t^2}{2n}+\frac{d(n)}{n}\right]^n=e^{t^2/2}=\text{ 標準常態分佈之 MGF，}$$

故由動差母函數之收斂定理可得 $Z_n \xrightarrow{d} Z \sim N(0,1)$，定理得證。(以上證明係以 Bain and Engelhardt 1992 一書 239 頁之證明為藍本加以潤飾而成)

為了讓讀者易於了解以上證明之整體架構，筆者將有關 CLT 之預備知識及其證明的輪廓繪製成流程圖如下：



最後，筆者要特別強調，以上證明乃針對"母體無限大，母體平均數、變異數、動差母函數皆存在，而樣本為 i.i.d."的假設下之 CLT 版本。至於其他假設條件下之 CLT 版本，則十分眾多，且其證明往往更為艱深，例如 Erdös and Rényi (1959)一文中所討論的 CLT，是屬於有限母體(finite population)的不置回抽樣(sampling without replacement)而產生"非 i.i.d.樣本"之情況下的版本(請參閱 Thompson 1992 一書 28 頁)，又如 Bowers *et al.* (1997)一書 39-40 頁談到"獨立但

非 i.d.樣本”之情況下的 CLT(應用於風險理論或精算學等領域)，等等，真是不勝枚舉。

3. 結語

統計學是國內大專院校許多科系的必修課程，也是令許多大專學生頭痛以及給授課教師們帶來困擾的科目。而中央極限定理之證明，因篏牽涉到大量的數理推導，更是令許多大專學生望之怯步，教師們也往往略而不教。本文針對中央極限定理之證明以及相關的預備知識做一個總整理，希望能對廣大的修習統計學之大專學生有所助益。

參考文獻

Bain, L. J. and Engelhardt, M. (1992). Introduction to Probability Theory and Mathematical Statistics, 2nd ed., Duxbury.

Belding, D. F. and Mitchell, K. J. (1991). Fundations of Analysis, Prentice-Hall.

Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., and Nesbitt, C. J. (1997). Actuarial Mathematics, 2nd ed., Society of Actuaries.

Chung, K. L. (1974). A Course in Probability Theory, 2nd ed., Academic Press.

Curtiss, J. H. (1942). A note on the theory of moment generating functions, *Annals of Mathematical Statistics*, **13**, 430-433.

Erdös,P. and Rényi, A. (1959). On a central limit theorem for samples from a finite population, *Publications of the Mathematical Institute of the Hungarian Academy of Science*, **4**, 49-61.

Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Vol. II, 2nd ed., Wiley.

Lukacs, E. (1970). Characteristic Functions, 2nd ed., Hafner, New York.

Rohatgi, V. K. (1976). An Introduction to Probability Theory and Mathematical Statistics, Wiley.

Thompson, S. K. (1992). Sampling, Wiley.

Widder, D. V. (1961). Advanced Calculus, 2nd ed., Prentice-Hall.