



# 行政院國家科學委員會專題研究計畫成果報告

## 雙重非中心化統計分配之計算和應用

計畫編號：NSC 96-2118-M-343-001

執行期限：96年8月1日至97年7月31日

主持人：陳仁義 教授 南華大學資訊管理系

### 一、中文摘要

雙重非中心化 beta 之統計量分配 (Doubly noncentral beta distributions) 在資料分析上扮演著重要的角色，其計算演算法的設計也是雙重非中心 F、t 分配的基礎，而隨機分配值(累積分配函數值)的計算法研究、分析和比較之文獻不多 (Chattamvelli, 1995; DasGupta, 1968; etc.)，在一般統計軟體中也沒有被完整程式化出來。在這個計畫中，我們利用常見的內外雙層無窮級數項表示式 (Johnson, Kotz, and Balakrishnan, 1995; etc.) 來作計算法之研究比較，並以 Benton and Krishnamoorthy (2003) 數值法為基礎，我們先將內層的 **singly noncentral beta distribution** 從兩個方向來改善之：一則是找到級數項所形成的無窮序列之真正峰點以作為數值疊代法的起始點，二來是更為有效地估計出有限級數項所形成的誤差界限。以此為基礎，我們利用了 Kocherlakota and Kocherlakota (1991) 的構想來完成雙層計算。此外，一些先前的研究成果 (Chen and Chou, 2000; Chen 2001; 2003; 2005) 也加以延伸和運用。所完成的三種數值方法之程式化、效率評估、分析和比較均已整理在 Chen (2008) 中。初步的結果顯示：三種數值演算法均可達到要求的精準度，而在一些情況 (e.g., noncentrality 參數值夠大時等)，我們所設計的方法具有最好的穩定性。

**關鍵詞：**非中心化 beta 分配、  
統計計算、誤差界限。

### Abstract

The doubly noncentral beta distribution plays an important role in data analysis. Its numerical computing methods are highly correlated to those of doubly noncentral F and t distributions. However, few studies and comparisons on their numerical algorithms have been done (Chattamvelli, 1995; DasGupta, 1968; etc.). In addition, they are not yet implemented in most of statistical software packages. In this project, the representation of doubly weighted series is used (Johnson, Kotz, and Balakrishnan, 1995; etc.). Based on these and methods of Benton and Krishnamoorthy (2003), we improve the algorithm for computing the inner part of **singly noncentral beta**. Three algorithms are developed and implemented in R. Our method can exactly find the mode of the infinite items in the inner part of series as well as more likely estimate a sharper error bound in finite steps. Subsequently, we use the idea of Kocherlakota and Kocherlakota (1991) to accomplish the computation of doubly weighted series. In addition, part of the results and the methods (Chen and Chou, 2000; Chen 2001; 2003; 2005) are applied and extended in this project. The three implemented numerical procedures performed under the general situations are competitive and have good achievements for their computational precisions. However, the larger the noncentrality values are, the greater the stability is for our proposed method. **Details are presented in Chen (2008).**

**Keywords:** Nocentral Beta Distribution;  
Statistical Computing;  
Error Bound.

## 二、緣由與目的

Doubly Noncentral beta 統計量分配的數學表示式和相關性質已經被推導出來 (Chattamvelli, 1995; DasGupta, 1968; Johnson, Kotz, and Balakrishnan, 1995; etc.)，其中的隨機分配之數學式子是以無限的雙層級數 (doubly weighted infinite series) 來表示，級數中的 incomplete beta integral 是為最重要份子之一，配合著其間雙權重函數 (double weight function) 的序列變化，正好可引用個人的先前研究成果 (Chen, 2001)。此外，非中心化分配的計算法中往往存在著一個重要數值計算問題，就是 noncentrality 參數值夠大時，常會造成數值近似誤植為零或者數值穩定性等問題，Doubly noncentral beta 的數值法也不例外，同樣發生了這類問題。Benton and Krishnamoorthy (2003) 等學者所提出的以權重函數 Mode 為中心點，同時地向正反兩個方向作數值疊代近似。在此計畫中，我們把這個概念應用到內層的權重函數之計算 (Chattamvelli, 1995, page 233)，形成了第一種方法 (Method I)；進一步地找到較為準確的級數項序列之真正峰點以作為數值疊代法的起始點，形成了我們設計的方法 (Method II)，且將近似的有限項之誤差估計加以改善；初始概念中的此類方式近似法 (Chattamvelli, 1995; Ding, 1995) 並沒有加入 Benton and Krishnamoorthy (2003) 的想法則形成了第三種方法 (Method III)，以作為我們評量比較的基準。基於這內層的 singly noncentral beta distribution 之結果，再接再力完成雙層的新計算法研究 (Kocherlakota and Kocherlakota, 1991)。同時運用個人初步研究成果 (Chen and Chou, 2000; Chen 2001; 2003; 2005) 且加以延伸，試圖改善參數值夠大時所造成的誤植為零或者數值精準和穩定性等問題。基於此所設計的一些情況，我們先行測試所程式化的數值結果，是否具有有一致性，進而作深入的分析 and 比較。

## 三、結果與討論

陳仁義 (Chen, 2001; 2003; 2005; Chen and Chou, 2000) 已經累積了統計量分配為 Singly Noncentral Distributions 的數值計算之經驗，主要考量是在誤差容忍度範圍之內，可以有效率而快速地找到其中有限項 (finite terms) 來求得滿足的近似值，同時需要考慮在計算過程當中應避免數值的 underflow/overflow 問題和其穩定性，此外，Benton and Krishnamoorthy (2003) 等研究者所提出的概念，更強化了這些問題的解決方案。以這些研究結果為基礎，我們推廣到雙權重函數情況的計算法研究，以三種的方式來近似 Doubly noncentral beta 之統計量分配的計算法，形成相應的數值法，並且以 R (R Development Core Team, 2008) 來完成程式化工作和效率評估。初步的研究結果整理在 **陳仁義 (Chen, 2008)**。主要的貢獻有 **一則** 配合著實際情況將一元的方法推展到二元的情況，並且簡化了參數的變化方式，**二來** 提高了計算法的精確和穩定性，避免非中心化參數值夠大時所造成的困擾問題，**三則** 以免費的 R 軟體來程式化了這些數值演算法，並且作了詳盡的分析和比較，使得應用者更為清晰的認識此統計量分配。

## 四、計畫成果自評

我們以統計科學理論為基礎，引用 Benton and Krishnamoorthy (2003) 所提出的概念，從多面向的數值近似法來計算 Doubly noncentral beta 統計量分配，這嘗試從一元的方法推展到二元的情況，難免碰上高維度推廣所會遭遇的諸多困難問題，初步的解決方案可以行得通，也表現得不錯，在後續和進行中的研究中，將會更為深入而持續強化這些數值近似之整合性。在學術研究方面，對雙重非中心化的統計量計算有較為廣泛而深入的探討，在統計與計算的跨領域研究有進一步的發展，在應用上也較為容易。參與人員在數值計算法和統計量的跨領域研究上有些基礎訓練，並且透過程式化工作中，培養出電腦化技能來。在跨領域的互動研究上已播下了種子而漸入佳境。

## 五、參考文獻

- [1] Benton, D. and Krishnamoorthy, K. (2003), "Computing discrete mixtures of continuous distributions: noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient," *Computational Statistics and Data Analysis*, **43**, 249-267.
- [2] Chattamvelli, R. (1995), "A note on the noncentral beta distribution function", *The American Statistician*, **49**, 231-234.
- [3] **Chen, Z.Y.** (2001), "On the Statistical Computation of the Sample Multiple Correlation Coefficient," *Journal of Statistical Computation and Simulation*, **70**, 299-324.
- [4] **Chen, Z.Y.** (2003), "Computing the Distribution of the Squared Sample Multiple Correlation Coefficient with S-System", *Communications in Statistics, Simulation and Computation*, **32**, No 3, 873-898
- [5] **Chen, Z.Y.** (2005), "The S-System Computation of Non-Central Gamma Distribution," *Journal of Statistical Computation and Simulation*, **75**, 813-829.
- [6] **Chen, Z.Y.** (2008), "On the Computation of Doubly Noncentral Beta Distributions," Technical Report, Department of Information management, Nan-Hua University.
- [7] **Chen, Z.Y.** and Chou, Y.C. (2000), "Computing the Noncentral Beta Distribution with S-System", *Computational Statistics and Data Analysis*, **33**, 343-360.
- [8] Ding, C. G. (1994). "On the computation of the noncentral beta distribution", *Computational Statistics and Data Analysis*, **18**, 449-455.
- [9] Frick, H. (1990), "A remark on Algorithm AS 226: computing noncentral beta probabilities", *Applied Statistics*, **39**, 311-312.
- [9] Johnson, N. J., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distributions*, Vol. 2, 2nd Ed., Wiley: New York.
- [10] Kocherlakota, K. and Kocherlakota, S. (1991), "On the doubly noncentral t distribution," *Communications in Statistics -- Simulation and Computation*, **20**, 23-31.
- [11] Lee, C. M. S. (1993), "On the computation of central and noncentral beta probabilities", *Journal of Statistical Computation and Simulation*, **43**, 1-10.
- [12] Lenth, R. V. (1987), "Algorithm AS 226: computing noncentral beta probabilities", *Applied Statistics*, **36**, 241-244.
- [13] Posten, H.O. (1993), "An effective algorithm for the noncentral beta distribution function", *The American Statistician*, **47**, 129-131.
- [14] R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.