

# 行政院國家科學委員會專題研究計畫 成果報告

## 非中心化 t 統計分配的計算法研究 研究成果報告(精簡版)

計畫類別：個別型  
計畫編號：NSC 95-2118-M-343-001-  
執行期間：95年08月01日至96年07月31日  
執行單位：南華大學資訊管理學研究所

計畫主持人：陳仁義

計畫參與人員：碩士班研究生-兼任助理：蔡孟哲、高于婷、趙堂鐘、鄭信源

處理方式：本計畫可公開查詢

中華民國 96 年 10 月 29 日



# 行政院國家科學委員會專題研究計畫成果報告

## 非中心化 t 統計分配的計算法研究

計畫編號：NSC 95—2118—M—343—001

執行期限：95 年 8 月 1 日至 96 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

非中心化 t 分配 (Doubly noncentral t distributions) 在資料分析上是很重要的統計量，其隨機性分配值 (機率密度函數值、累積分配函數值、百分位數值等) 計算法尚未被廣泛研究、分析和比較，在一般的統計軟體中也沒有被完整程式化出來。在這個研究計畫中，我們是以 Benton and Krishnamoorthy (2003) 數值計算法為基礎，從兩個方向來近似之：一則是間接地利用單層的 noncentral t distribution 之結果以完成雙層的計算 (Kocherlakota and Kocherlakota, 1991)；二來是直接地完成雙層表示式 (Bulgren, and Amos, 1968; Krishnan, 1968; etc.) 之新計算法研究。為達成所要求精確度，我們運用先前研究成果 (Chen and Chou, 2000; Chen 2001; 2003; 2005) 且加以延伸，詳細的數值演算法之程式化、效率評估、研究結果分析和比較均已整理在 Chen (2007) 中，初步結果顯示：所提出的兩種數值方法互有勝出，整體而言，兩者優於傳統近似法 (Broda and Paolella, 2007; Bulgren, 1974; Bulgren and Amos, 1968; etc.)，且可避免 noncentrality 參數值夠大時，其數值近似值會誤植為零或計算速度變得很慢等問題出現。

**關鍵詞：**非中心化 t 分配、  
統計計算、誤差界限

### Abstract

The doubly noncentral t distribution have been received much attention (Broda. and Paolella, 2007; Bulgren and Amos, 1968; Johnson, Kotz and Balakrishnan, 1995; Kocherlakota and Kocherlakota, 1991; Krishnan, 1967; 1968; etc.). Its form of series representation is in terms of doublely infinite sum of Poisson weights of the incomplete beta function. The derived procedures for numerical computation differ much or less on the error bounds, the simplicity of representations, and the approximation methods (Johnson, Kotz and Balakrishnan, 1995; etc.). Most of them are for computing cumulative probabilities. However, the larger the noncentrality values are, the greater the variations of their evaluation accuracy and time efficiency. In this project, we investigate the numerical algorithms for cdf approximations using the methods of Benton and Krishnamoorthy (2003) and Chen (2001). Based on these, two numerical procedure are proposed: one representing the double seies involving singly noncentral t by Kocherlakota and Kocherlakota (1991); the other forming with double summationn (Bulgren, and Amos, 1968; Krishnan, 1968; etc.). Their performance of the two proposed methods are competitive. Details are presented in Chen (2007). However, their results outperform the ad hoc calculations or approximations (Broda and Paolella, 2007; Bulgren, 1974; Bulgren and Amos; etc. ) Morevoer, it improves the time efficiency and avoids the trouble problem when noncentrality parameter values are large enough.

**Keywords:** Incomplete Beta Integral;  
Statistical Computing;  
Error Bound; .

## 二、緣由與目的

Doubly Noncentral  $t$  統計量分配的數學表示式和相關性質已被推導出來 (Bulgren, 1971; Bulgren and Amos, 1968; Kocherlakota and Kocherlakota, 1991; Krishnan, 1967; 1968; etc.)，其中隨機分配的數學式是以無限雙層級數(doubly infinite series)來表示，級數中的 incomplete beta integral 為最重要份子之一，配合著雙權重函數 (doubly weight function) 的序列變化，正好引用個人的先前研究成果(Chen, 2001)。另一方面，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計算速度變慢等問題，Doubly Noncentral  $t$  的數值法也不例外，同樣發生了這類問題。Benton and Krishnamoorthy (2003)等學者提出以權重函數 Mode 為中心點，同時向正反兩個方向作數值疊代近似。在此計畫中，我們把這個概念推廣到雙權重函數情況，運用此函數的二元化中心點，以兩種方式來近似：一則是間接地利用單層的 noncentral  $t$  distribution 之結果來完成雙層的計算 (Kocherlakota and Kocherlakota, 1991)；二來是直接地完成雙層表示式 (Bulgren, 1974; Bulgren and Amos, 1968; Krishnan, 1968; etc.)之新計算法研究。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計算速度變慢等上述問題。

## 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 問題和計算速度，此外，Benton and Krishnamoorthy (2003) 和 Chattamvelli and Shanmugam (1997) 等研究學者所提出的概念，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二元化中心點，以兩種方式來近似 Doubly Noncentral  $t$  統計量分配的計算法，形成了相應的數值演算法，並且以 R (R Development Core Team, 2005) 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, 2007)，其中顯示，所提出的兩種數值方法互有勝出，並且我們已作了相當程度改善，整體而言，兩者均優於傳統近似法 (Broda and Paolella, 2007; Bulgren, 1974; Bulgren and Amos, 1968; Mudholkar and Chaubey, 1976; etc.)。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**提高了計算法的效率和精確度，避免非中心化參數值夠大時所造成的困擾問題，**三則**電腦化這些計算方法則更為容易而有效率，應用此統計量分配較為方便。

## 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  $t$  統計量分配，這個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，在後續和進行中的計畫『雙重非中心化統計分配之計算法與應用』，將會更為深入而持續地強化和改善這些數值近似法。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] Bulgren, W.G. (1974), "Probability integral of doubly noncentral t-distribution with degrees of freedom  $n$  and non-centrality parameters  $\delta$  and  $\lambda$ ," in *Selected Tables in Mathematical Statistics*, edited by Harter and Owen, Providence RI: American Mathematical Society, 1-138.
- [3] Bulgren, W.G. and Amos, D. E. (1968), "A note on representations of doubly noncentral t-distribution," *Journal of the American Statistical Association*, **63**, 1013-19.
- [4] Broda, S. and Paoletta, M.S. (2007), "Saddlepoint approximations for the doubly noncentral t distribution," *Computational Statistics and Data Analysis*, **51**, 2907-18.
- [5] Chattamvelli, R. and Shanmugam, R. (1997), "Computing the noncentral beta distribution function," *Applied Statistics*, **46**, 146-156
- [6] **Chen, Z.Y.** (2001), "On the Statistical Computation of the Sample Multiple Correlation Coefficient," *Journal of Statistical Computation and Simulation*, **70**, 299-324.
- [7] **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
- [8] **Chen, Z.Y.** (2005), "The S-System Computation of Non-Central Gamma Distribution," *Journal of Statistical Computation and Simulation*, **75**, 813-829.
- [9] **Chen, Z.Y.** (2007), "On the Computation of Doubly Noncentral t Distributions," Technical Report, Department of Information management, Nan-Hua University.
- [10] **Chen, Z.Y.** and Chou, Y.C. (2000), "Computing the Noncentral Beta Distribution with S-System," *Computational Statistics and Data Analysis*, **33**, 343-360.
- [11] Johnson, N. J., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distributions*, Vol. 2, 2nd Ed., Wiley: New York.
- [12] Kocherlakota, K. and Kocherlakota, S. (1991), "On the doubly noncentral t distribution," *Communications in Statistics -- Simulation and Computation*, **20**, 23-31.
- [13] Krishnan, M. (1967), "The moments of a doubly noncentral t-distribution," *Journal of the American Statistical Association*, **62**, 278-287.
- [14] Krishnan, M. (1968), "Series representations of the doubly noncentral t-distribution," *Journal of the American Statistical Association*, **63**, 1004-1012.
- [15] Mudholkar, G.S. and Chaubey, Y.P. (1976), "A simple approximation for the doubly noncentral t distribution," *Communications in Statistics, Simulation and Computation*, **B5**, No 2&3, 85-92.
- [16] R Development Core Team (2005). 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>.