

國立政治大學統計學系
碩士學位論文

KLIC作為傾向分數配對平衡診斷之可行性探討

**Using Kullback-Leibler Information Criterion on
balancing diagnostics for baseline covariates between
treatment groups in propensity-score matched samples**

指導教授：江振東 博士

研究生：李珮嘉 撰

中華民國一百零四年二月

摘要

觀察性研究資料中，透過傾向分數的使用，可以使基準變數在實驗與對照兩組間達到某種程度的平衡，並可視同為一隨機試驗，進而進行有效的統計推論。文獻中有關平衡與否的診斷，大多聚焦於平均數與變異數的比較。本文中我們提出使用 KLIC(Kullback-Leibler Information Criterion)及 KS(Kolmogorov and Simonov)兩種比較分配函數差異的統計量，作為另一種平衡診斷工具的構想，並針對其可行性進行探討與評比。此外，數據顯示 KLIC 及 KS 與透過傾向分數配對的成功比例呈現負相關。由於配對成功比例過低將導致後續統計推論結果的侷限性，因此本文也就 KLIC 及 KS 作為是否進行配對的一個先行指標之可行性作探討。模擬結果顯示，二者的答案均是肯定的。

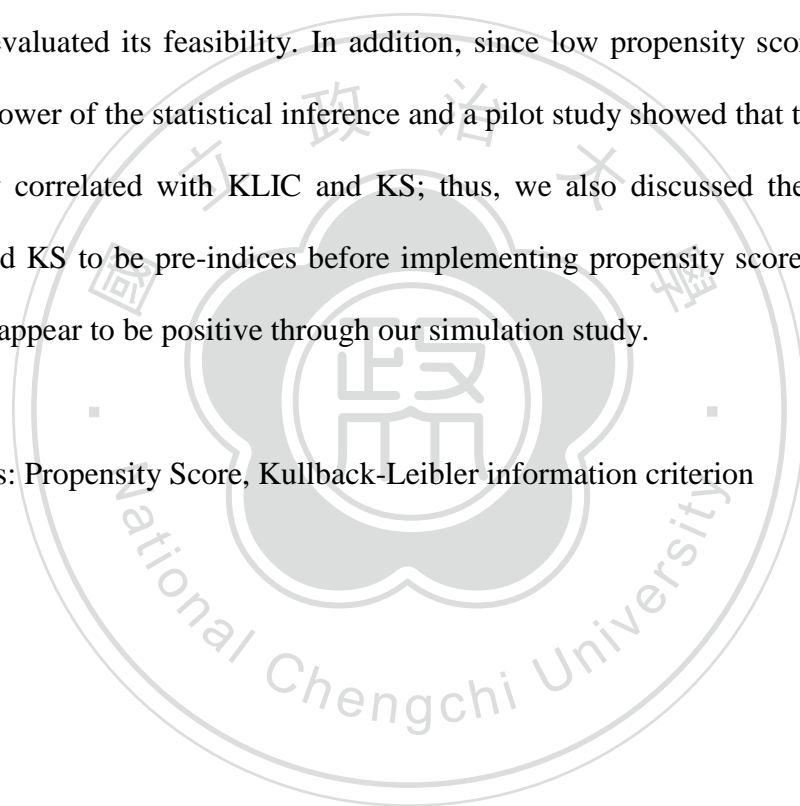
關鍵字：傾向分數、Kullback-Leibler information criterion



Abstract

In observational studies, propensity scores are frequently used as tools to balance the distribution of baseline covariates between treated and untreated groups to some extent so that the data could be treated as if they were from a randomized controlled trial (RCT) and causal inferences could thus be made. In the past, balance or not was usually diagnosed in terms of the means and/or the variances. In this study, we proposed using either Kullback-Leibler Information Criterion (KLIC) or Kolmogorov and Simonov (KS) statistic as a diagnostic measure, and evaluated its feasibility. In addition, since low propensity score matching rate decreases the power of the statistical inference and a pilot study showed that the matching rate was negatively correlated with KLIC and KS; thus, we also discussed the possibilities of using KLIC and KS to be pre-indices before implementing propensity score matching. Both considerations appear to be positive through our simulation study.

Key words: Propensity Score, Kullback-Leibler information criterion



目錄

第一章 緒論.....	1
第二章 文獻探討.....	3
第一節 KLIC.....	3
壹、Shannon Entropy.....	3
貳、Kullback-Leibler Information Criterion.....	4
第二節 傾向分數.....	5
第三章 研究方法與設計.....	7
第一節 研究目的.....	7
第二節 研究設計.....	8
壹、臨界值的求取.....	8
貳、檢定力比較.....	10
參、樣本數比例對照.....	13
第四章 研究結果與討論.....	16
第一節 臨界值的選定.....	16
第二節 型一錯誤機率及檢定力.....	18
第三節 配對前後對照表.....	24
第五章 實證分析.....	28
第一節 主題分析.....	28
第二節 資料來源與變數定義.....	28
壹、資料來源.....	28
貳、變數說明.....	29
第三節 分析結果.....	29
壹、分析步驟.....	29

貳、分析結果.....	30
第六章 結論與建議.....	32
第一節 結論.....	32
第二節 未來研究方向建議.....	33
參考文獻.....	34
附錄.....	35



第一章 緒論

在許多社會科學的研究中，大多使用觀察性的研究方式來收集資料，相較於自然科學的研究，這通常是不得不的作法，也是唯一的選擇。然而這種資料收集的方式最為人詬病的，莫過於實驗組與對照組在某些基準變項上，可能呈現出明顯的差異，導致後續統計推論結果的偏誤。針對此一缺憾，Rosenbaum and Rubin(1983)提出傾向分數(propensity score)的構想，並證明傾向分數其實是一種平衡分數(balancing score)，為觀察性研究的資料分析，開啟一道可行之路[1]。Austin(2011)指出，傾向分數的使用可以讓不同組別間無法隨機控制的變量獲得某種程度的平衡，減少干擾因子的影響並達到隨機化配置的效果，借以提供較高可信度的推斷結論[2]。

現行文獻中有關實驗、對照組間平衡與否的診斷，多半聚焦於兩者間相關變項平均值與變異數的比較。相較之下，兩者間分配函數比較的建議，較不多見。Frölich(2004)將Kullback-Leibler 訊息指標(Kullback-Leibler Information Criterion, KLIC)和傾向分數做一個連結，用以判斷配對前的兩實驗組別的分配差異[3]。Busso, DiNardo, McCrary, et al.(2011)就兩組分配不同重疊比例下傾向分數的好壞表現作延伸討論[4]。由於欲比較兩分配間差異時，Kolomgorov and Simonov(KS)檢定統計量是一般常被使用的工具，其表現在大多數情況下具有一定效果。因此本文的主要目的之一，就是要針對 KLIC 及 KS 作為一項平衡診斷工具的可行性，進行評比與探討。

再者，文獻顯示透過傾向分數配對過後的資料多半都可以達到平衡的效果，然而如果配對成功樣本太少，即使配對後非常均衡，則統計資料分析過後的結論，其推論範圍恐怕會有侷限性。因此本文的另一個目標則是想要了解「KLIC 與重疊比例的可能關係」，也就是說兩者分配差異大小與配對成功比例數是否有關？關係又如何？

本文共分為六章，第二章為文獻探討，將針對會使用到的方法做一個簡單的回顧；第三章為研究方法，就主要目的敘述並設計模擬架構；第四章為研究結果，會針對模擬出的資料與圖表進行討論與說明；第五章為實證分析，將本文所設計的方法應用在實務

資料上，評估可行性；最後則是總結，針對文中所發現的結果與限制做一個統整，並檢討未來可行的方向。



第二章 文獻探討

此章將就 Kullback-Leibler Information Criterion(KLIC)及傾向分數(propensity score)這兩項後續章節中所會涉及的議題，進行簡單回顧。

第一節 KLIC

壹、 Shannon Entropy

Entropy 一詞最早在熱力學由 Boltzman 於 1872 年提出，後來 Shannon 在 1948 年延伸到訊息理論(information theory)上並正式定義，詳述如下[5]。

已知一離散隨機變數 $X = \{x_1, \dots, x_n\}$ ，其機率密度函數(probability mass function)為 $f(X)$ ，定義 $H(X)$ 為

$$\text{Shannon Entropy} := H(X) = - \sum_{i=1}^n f(x_i) \ln f(x_i) \stackrel{\text{def}}{=} H(F)$$

藉由上述定義可以延伸許多性質如下：

- (1) 恆為正數： $H(X) \geq 0, \forall x_i \in X$ 。
- (2) 平均訊息指標：將原式作延伸可得 $H(X) = - \sum_{i=1}^n \ln f(x_i) f(x_i) = E_X[-\ln f(X)]$ ，因此 $H(X)$ 可視為一個期望值的概念，表示平均而言此機率函數所提供的訊息。
- (3) 不確定性指標：我們知道當 x_i 的發生機率越低，表示不確定性越高。由定義對照可知，此時其 $f(x_i)$ 會近似於 0，使得 $-\ln f(x_i) \rightarrow \infty$ 。另外若我們知道 x_i 確定發生或不發生的情況，則無不確定性可言。由於當 x_i 不發生時， $f(x_i) = 0$ ； x_i 確定發生時，其機率 $f(x_i) = 1$ ，這兩者對 $H(X)$ 所造成的貢獻都是 0，所以 $H(X)$ 可解釋為用於測量隨機變數 X 的不確定性。

另外，Ullah(1996)也提到由於影響 $H(X)$ 大小決定於其機率函數 $f(X)$ 而非變數值本身，故文獻上大多使用 $H(F)$ 以作為強調[6]。

儘管 Shannon entropy 有良好的定義，也因為意義簡單明白被廣泛應用，但生活上許多例子大多建構在連續狀態下，所以將 $H(F)$ 中的總和改為積分符號做為調整，卻因此產生了另外的問題：「若隨機變數為連續，其機率密度函數的選取可能會使 entropy 產生負值」，故延生出其他處理方法，KLIC 為其中一種。

貳、 Kullback-Leibler Information Criterion

一、起源

Kullback and Leibler(1951)將 Shannon entropy 做一延伸，用以測量兩機率測度的差異(discrimination)，後來稱之為 KLIC(也稱 relative entropy)，其定義如下[7]：

$$KLIC := I^{KL}(F:G) = \int f(x) \ln \frac{f(x)}{g(x)} dx = E_F \left[\ln \frac{f(x)}{g(x)} \right]$$

由上式可看到幾點特性：

(1) 恆為正數：利用詹森不等式¹(Jensen's inequality)可輕易證明如下。

$$I^{KL}(F:G) = - \int f(x) \ln \frac{g(x)}{f(x)} dx \leq - \ln \left(\int f(x) \frac{g(x)}{f(x)} dx \right) = - \ln(1) = 0$$

(2) 等號成立表示測度相等：亦即 $I^{KL}(F:G) = 0 \Leftrightarrow f(x) = g(x)$ 。

(3) 無對稱性(non-symmetry)：由定義明顯可得，故 Kullback and Leibler(1951)也定義了 Kullback-Leibler divergence: $J^{KL}(F:G) = I^{KL}(F:G) + I^{KL}(G:F)$ 以滿足其對稱性。此概念最先由 Jeffreys(1946)提出，故也稱 Jeffreys information。

(4) 不符合三角不等式：Ullah(1996)整理了許多不同距離及其衍生，其中指出 KLIC 並不是一個距離(metric)概念，因為不符合對稱性與三角不等式，即便是 Jeffreys information 也仍不滿足距離公設，故也不為距離。

¹ 詹森不等式，若 f 為一凸(convex)函數，則滿足 $f(\int g(x)dx) \leq \int f(g(x))dx$ 。

二、應用

由於 KLIC 可以做為檢視兩測度訊息差異的指標，後來被廣泛使用在模型選取，當應用於實務時，若將 f 視為所真實的模型， g 為用以估計的模型，兩者皆為機率測度，應滿足下列條件：

$$(1) \int f(x)dx = \int g(x)dx = 1$$

$$(2) g(x) = 0 \implies f(x) = 0$$

則 KLIC 在此將兩測度的差異解釋為：「使用模型 g 來估計背後真實 f 所遺失的訊息」，然而實務的困難點是，我們永遠不知道真實 f 為何，因此將 KLIC 定義展開如下

$$\int f(x) \ln f(x) dx - \int f(x) \ln g(x) dx$$

這邊的重點在於前項表示真實的分配故無法測量，雖然如此，這真實分配卻是固定的，代表前項可視為一個"常數"，而後項(後來將此稱為 cross entropy)可以經由樣本模型等等估計，因此欲選取最佳的模型，就表示希望遺失的訊息最少，即最小的 KLIC，也意味著最小的 cross entropy，故實務上會去比較各個模型的 cross entropy 以求得最佳模型。另外許多文獻上應用可發現，cross entropy 必須經由比較才有意義，故而延伸出許多估計方法，如 AIC、BIC、QAIC 等等。

第二節 傾向分數

在實驗性研究(experimental study)中，要判斷特定變數的效果時，我們可以利用隨機實驗的方式來控制其他變數以達成隨機配置的目的，進而推斷是否具有因果關係；然而生活中絕大部分是屬於觀察性研究(observational study)，其所觀察到的現象不見得只受到特定變數的影響，可能是來自其他的變項交互影響而成，也因此推斷關係上造成困擾與錯誤估計，故近年來發展出傾向分數(propensity score)來處理類似問題。

Rosenbaum and Rubin(1983)提出傾向分數一詞，將一個均衡分數(balancing score)定義為傾向分數 $e(X)$ ：「給定各控制變量(covariates)的條件下其受測者所指派到特定處理(treatment)組的機率。」給定傾向分數下，其組別的分類應與基準變量獨立。若以數學式表示則為

$$e(X) := P(Z|X_1, X_2, \dots, X_n)$$

其中

$$\begin{cases} X_i : \text{各控制變量} \\ Z : \text{實驗指派變數} \\ X_i \perp Z | e(x) \end{cases}$$

若將傾向分數的性質做一個意義上的延伸，表示在相同的傾向分數下，其不同實驗組別的其他控制變項分配結構相近，可視為一個模擬的隨機試驗(randomization experiment)，並據以進行資料分析。

傾向分數常透過羅吉斯模型來估計，後續文獻中也陸續發展出四種常用之傾向分數使用方法，(1)配對(matching)、(2)分層(stratification)、(3)加權調整(inverse probability of treatment weighting, IPTW)、(4)變量調整(covariate adjustment)，其中最被廣泛使用的是配對方法。由於真實的傾向分數為一個均衡分數，故事後兩組之間的平衡診斷(balance diagnostics)是必要的。常用的診斷工具包含標準化差異(standardized difference)、配對前後變異數比(variance ratio)、盒鬚圖(boxplot)、Q-Q plot、機率密度函數(PDF)、累積機率函數(CDF)等其他圖形比較法(參見 Austin(2009))[8]。

第三章 研究方法與設計

本章將就研究目的與模擬實驗設計的細節部分做說明。

第一節 研究目的

文獻中有關傾向分數平衡診斷的探討，主要著眼於配對後的資料是否呈現一致的現象。常見的診斷工具如標準化差異(standardized difference)、變異數比率(variance ratio)、勝算比(odds ratio)等等，較少有涉及到分配函數比較的論述，以及配對數比例的探討。

而以往比較兩分配差異的方式最常使用 KS 雙樣本檢定(two-sample Kolmogorov-Smirnov test)，不同於 KLIC，KS 檢定的統計量為兩組之間累積機率函數的最大差異，其優點是不受樣本大小與分配限制，使用方式十分簡便且也因具有一定的檢定力，故被廣泛的使用，但做為評估整體的分配差異的一個指標似乎不如 KLIC 來的具有一般性，因此本文主要想要了解使用 KLIC 作為平衡診斷工具的可行性，並與 KS 統計量進行比較。此外我們認為過少比例的配對數在實用上可能較不充分，因此也值得探討。Frölich(2004)除指出越高的 KLIC 值顯示出對照—實驗配對兩組之間的差異越大外，也就兩組之配對個數比與傾向分數的 KLIC 值之間的相關性作了些許探討，故我們想要進一步延伸：「KLIC 值所顯示的差異是否與配對數變化有關？」，進而去考慮是否值得做傾向分數配對。由於傾向分數配對所需的時間會隨樣本數增加，尤其在處理巨量資料時特別耗時，而 KLIC 的計算所需的時間是大約是配對的一半，故或許可以在配對前當作一個是否進行配對的先行指標。

本文主要著重在 KLIC 與 KS 的比較。首先，我們將分別針對對稱、左偏、右偏等三種不同偏態的分配來進行探討。在虛無假設成立的前提下，也就是實驗、對照兩分配具有相同分配，我們將在不同樣本數的情況下，先找出 KS 與 KLIC 統計量兩者的 95% 百分位數，作為後續進行假設檢定時的臨界值(critical value)。其次，我們將依據三種分配的不同參數變化—位置(location)、散佈(scale)、形狀(shape)來比較 KLIC 與 KS 兩者之間檢定力的差異。最後，我們將依不同相關強弱產生兩變量的傾向分數，做配對前後的

實驗組樣本數變化比，並與 KLIC 與 KS 配對前對照關係，再以配對後之 KLIC 與 KS 統計量當作一個判斷的依據，最後計算 KLIC 與配對各別所需的時間。

第二節 研究設計

KLIC 又稱 relative entropy(Ullah,1996)，儘管理論上其值永遠為一正數，但若以兩樣本實際計算過程中可能因定義域不同產生正負相消的問題，且故這邊除了原始的三種 KLIC 定義之外，各別加了絕對值後去比較，分別定義如下：

$$(1) KLIC1 = \int \left(\ln \frac{p(x)}{q(x)} \right) p(x) dx = \int p(x) (\ln p(x) - \ln q(x)) dx$$

$$(2) KLIC2 = \int \left(\ln \frac{q(x)}{p(x)} \right) q(x) dx = \int q(x) (\ln q(x) - \ln p(x)) dx$$

$$(3) KLIC3 = \frac{1}{2} \left(\int \left(\ln \frac{p(x)}{q(x)} \right) p(x) dx + \int \left(\ln \frac{q(x)}{p(x)} \right) q(x) dx \right) = \frac{1}{2} (KLIC1 + KLIC2)$$

$$(4) KLIC4 = \int \left| \ln \frac{p(x)}{q(x)} \right| p(x) dx = \int p(x) |\ln p(x) - \ln q(x)| dx$$

$$(5) KLIC5 = \int \left| \ln \frac{q(x)}{p(x)} \right| q(x) dx = \int q(x) |\ln q(x) - \ln p(x)| dx$$

$$(6) KLIC6 = \frac{1}{2} \left(\int \left| \ln \frac{p(x)}{q(x)} \right| p(x) dx + \int \left| \ln \frac{q(x)}{p(x)} \right| q(x) dx \right) = \frac{1}{2} (KLIC4 + KLIC5)$$

壹、臨界值的求取

我們考慮三種偏態的分配，依序分別為對稱分配、左偏分配、右偏分配，並透過 R 軟體來進行模擬。每一次的模擬實驗皆從相同分配中生成兩組樣本去計算相同分配前提下的 KLIC 與 KS 值，藉由重複實驗 10000 次來找出 95% 的臨界值。樣本數的部份，我們考慮了 1000、500、200 這三種情況，來比較可能的變動差異。三種分配詳述如下：

(1) 對稱分配：常態(Normal)分配

$$N(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \mu \in R, \sigma > 0$$

(2) 左偏分配：偏斜常態(Skew-normal)分配

$$SN(x|\mu, \sigma, \alpha) = \frac{1}{\pi\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \int_{-\infty}^{\alpha\left(\frac{x-\mu}{\sigma}\right)} e^{-\frac{t^2}{2}} dt \quad \mu, \alpha \in R, \sigma > 0$$

(3) 右偏分配：Fréchet分配

$$Fréchet(x|\mu, \sigma, \alpha) = \frac{\alpha}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^{-1-\alpha} \quad \mu \in R, \sigma, \alpha > 0$$

圖 3-2-1 為三種分配的機率密度函數分布圖的比較。

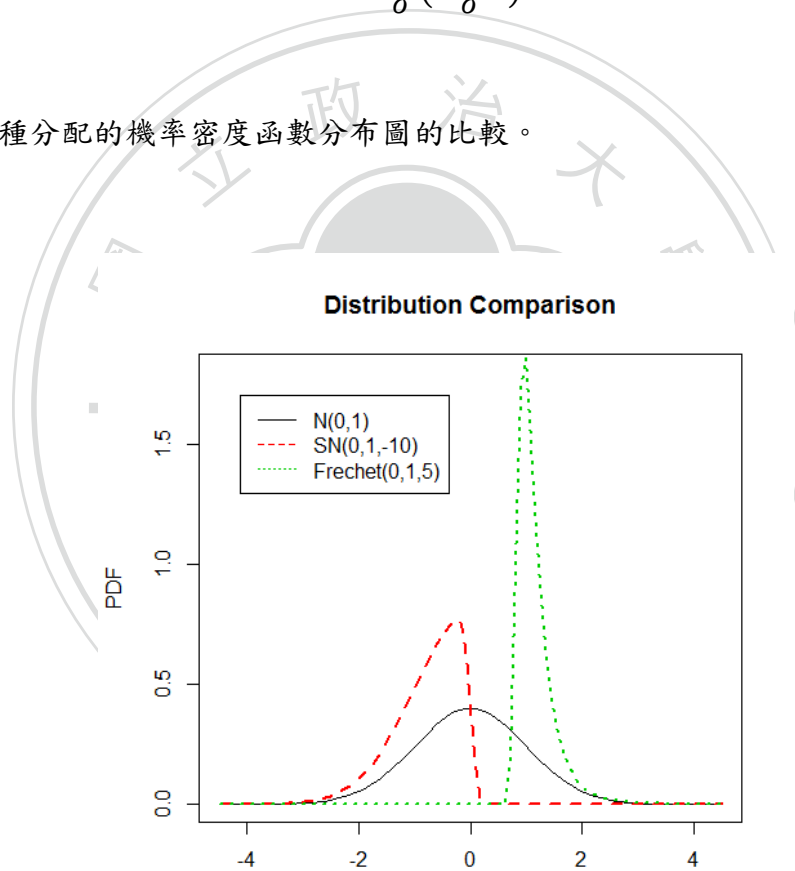


圖 3-2-1 三偏態 PDF 比較

貳、檢定力比較

依據前述說明求出 95% ($\alpha = 0.05$) 臨界值作為判斷實驗、對照兩分配是否相同的準則後，我們將針對三類分配的三種參數的不同變化：(1)位置—平均數移動、(2)散佈—變異數大小、(3)形狀—厚薄尾差異等，並比照三種樣本數大小(1000、500、200)，來探討各個統計量檢定力的強弱。檢定力的計算仍是架構在 10000 次模擬所得出的結果。

在第一類對稱分配中，以常態分配的平均數與標準差作為位置與散佈參數的對照，但由於常態分布並無厚尾的參數控制，這邊以同為對稱分配但具厚尾特性的 t 分布變動自由度代替比較。由於理論上的左偏例子不多，第二類左偏分配選用偏斜常態(SN)分配，是用常態去傾斜，所以最極端的情形只有到半常態分配(half-normal)；由第壹部分機率函數可知其 α 是控制常態偏斜的重要參數，也屬於形狀參數的一種，用以控制偏斜的程度，由圖 3-2-2.1 可知 α 的正負會決定偏態的方向， $\alpha = 0$ 時為對稱常態分配， $\alpha > 0$ 表示右偏， $\alpha < 0$ 為左偏，依次分別為紅色虛線、黑色實線、綠色虛線；接著去測驗 α 不同的負值所產生的變化，如圖 3-2-2.2 所示可看出 α 從-2到-10，偏態逐漸明顯，因此我們折衷選擇適當的偏斜程度，以 $\alpha = -10$ 作為我們模擬的基準分配；儘管 α 為形狀參數，圖 3-2-2.2 也觀察到不論 α 的改變為何並不會對尾部的分布造成影響，故我們試著考慮 Azzalini (2013, p.101-p.104)介紹的偏斜常態所延伸之偏斜 t 分配(skew- t distribution)來控制厚尾分布作為薄厚尾對照的分配[9]，如圖 3-2-2.2 紫色虛線所示，相較於黑色實線的偏斜常態分配，兩者除了尾部的變動之外，在其他地方並無明顯差異，故最後以偏斜 t 分配當作形狀參數的對照。

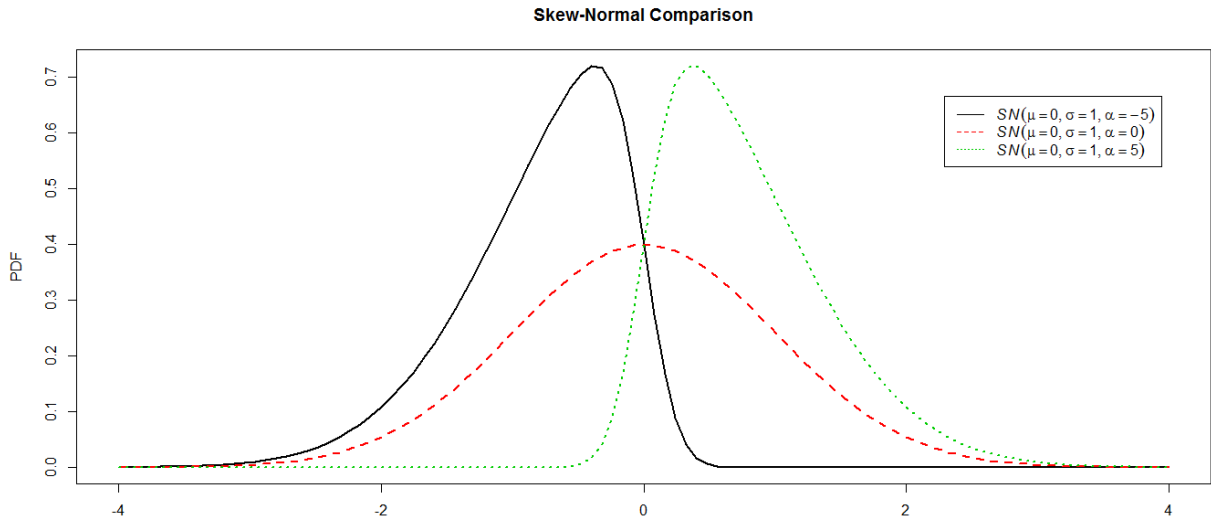


圖 3-2-2.1 不同 α 正負變化所影響之偏斜常態 PDF 對照

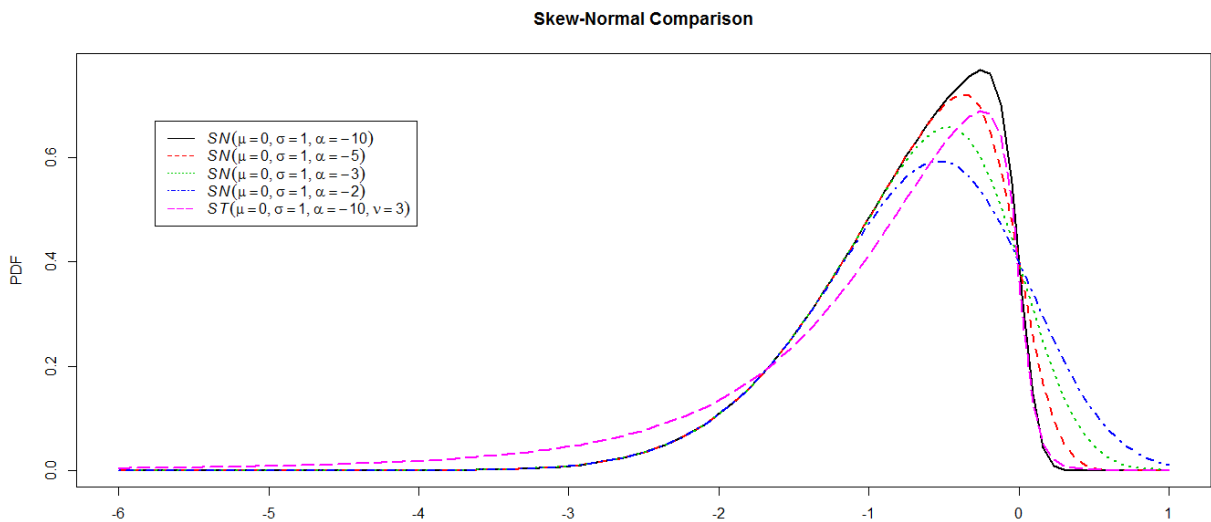


圖 3-2-2.2 偏斜常態與偏斜 t 分配對不同形狀參數變化之 PDF

因此總結左偏分配的選取，決定除了在偏態係數上選負號之外，其他位置與變異參數皆和常態分配一樣，以平均數和標準差變化控制，也因偏斜常態不具有控制薄厚尾表現的形狀參數，故以偏斜 t 分配來控制厚尾分布。實務上有許多常見的右尾分配，因Fréchet分配函數直接取決於三類的參數，故以此為第三類分配模型的選取，相較於左偏採用的偏斜常態，Fréchet本身具有可變動的形狀參數，變動不同參數對應的機率圖如圖 3-2-2.3 所示，其中黑色實線皆代表各個偏態的預設分配，不同虛線則是在固定位

置參數(μ)下，分別變動散佈參數(σ)與形狀參數($\nu; \alpha$)的情形。由圖 3-2-2.3(左)可看到在其他條件不變下，整體變異隨散佈參數(σ)增加而變大，機率密度函數(PDF)的變動由黑色實線到紅色虛線進而到綠色點線；而比較偏態 t 分配則可以發現隨著形狀參數(ν)下降，其尾部分布較原始的預設分配來的厚一些，PDF 變化由黑色實線到藍色虛線再到紫色虛線。同樣地，由圖 3-2-2.3(右)可比照在其他條件不變之下，散佈參數的增加會使整體變異上升，造成 PDF 較為分散，依序如圖中黑色實線、紅色虛線、綠色點線；而形狀參數(ν)的下降也使 PDF 呈現較厚尾的現象，變化由黑色實線、藍色虛線到紫色虛線。相較於左偏的偏斜常態分配而言，Fréchet不論在變異或形狀的變動上皆比偏斜常態來的敏感。

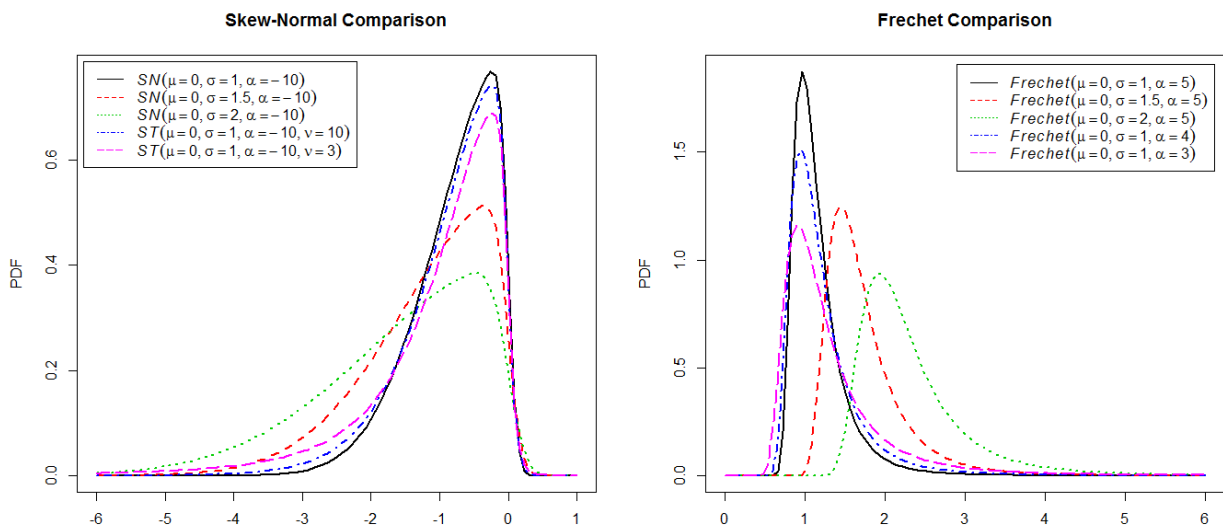


圖 3-2-2.3 左偏(圖左)與右偏(圖右)在不同參數下的 PDF

我們將依不同參數值來生成各種不同的對立假設分配，並與相對應的虛無假設分配— $N(0,1)$ 、 $SN(0,1,-10)$ 、 $Fréchet(0,1,5)$ —求出 KLIC 值與 KS 統計量後，進行檢定力的估算。假設檢定步驟如下：

(1) 建立假說： $\begin{cases} H_0 = \text{兩樣本來自相同母體(沒有差異)} \\ H_a = \text{兩樣本來自不同母體(有差異)} \end{cases}$ 。

(2) 比較臨界值：理論上，兩樣本分配結構相同時，KLIC 與 KS 統計量皆為 0；而當分配差異越大時，兩者皆隨之遞增，故以超過(大於)各自的 95% 臨界值視為拒絕兩分配相等的依據。

(3) 信心水準與檢定力：將上述之拒絕次數換算為此 10000 次模擬的比例，分為正確拒絕 H_0 與錯誤拒絕 H_0 兩種情形，各代表

a. $P(\text{拒絕 } H_0 | H_0 \text{ 為真}) = \alpha = \text{信心水準}$

b. $P(\text{拒絕 } H_0 | H_a \text{ 為真}) = 1 - \beta = \text{檢定力}$ 。

參、樣本數比例對照

我們猜測藉由傾向分數來進行配對，配對後所得到的配對數的多寡應該與原始資料中實驗組與對照組的分配結構有關。架構在這個想法之下，我們以為利用配對前實驗組與對照組的觀察值所計算得到的 KLIC 值，或許可以做為預估配對比率高低的一個先行指標。如此一來，如果預測出來的配對比率不是很高的話，或許就不需要繼續進行效果可能不大的配對步驟了。

為了瞭解配對比率是否與資料的原始分配有關，首先我們採用 Austin(2007)的建議，藉由不同的迴歸係數($\ln 2$ 、 $\ln 5$ 、 $\ln 10$)來隨機生成實驗組與對照組樣本[10]。圖 3-2-3 為總樣本數為 2000 情況下，模擬樣本傾向分數的機率密度函數(PDF)圖，由圖左至右係數分別為 $\ln 2$ 、 $\ln 5$ 、 $\ln 10$ ；黑色實線為實驗組，綠色虛線為對照組。上排為配對前的情況，可以看到兩者傾向分數的 PDF 由左而右差異漸大；而在配對之後(如下排所示)可以看到兩組之間分配幾乎一模一樣，顯示出非常良好的平衡結構。由於配對是以實驗組為主，因此配對後的分配較近似於原始實驗組的結構靠近，不過相較於原始分配仍有些許差異，尤其在兩樣本分配原先落差就很大的情況，差異會更加明顯。

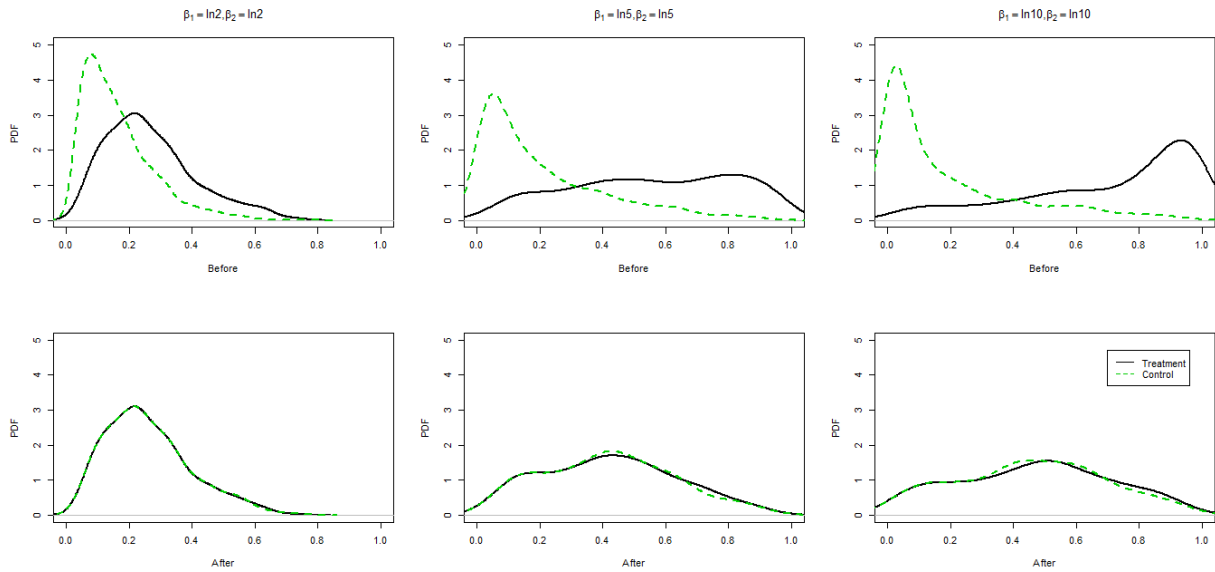


圖 3-2-3 樣本數 2000 之模擬資料中，配對前後傾向分數 PDF 的變化

接著我們試著觀察此三類樣本配對比率的差異，發現隨著原始兩組的差異增加，配對比率會明顯下降，甚至減少幾乎一半以上，如表 3-2-3 最右一行所示，實驗組原先的 761 筆樣本中成功配對者只有 330 筆。由於配對樣本數相對而言過少，若據此做推論，此結果可能有失充分。

表 3-2-3 樣本數 2000 之各係數對應樣本數前後比

係數	$\beta_1 = \beta_2 = \ln 2$	$\beta_1 = \beta_2 = \ln 5$	$\beta_1 = \beta_2 = \ln 10$
前	365	630	761
後	360	420	330
配對比例	0.9863	0.6667	0.4336

因此這一部分的模擬實驗中，我們希望可以藉由配對前後 KLIC 值以及 KS 統計量與配對樣本比率來探討期間可能的相關性。資料生成步驟如下：

- (1) 隨機生成兩變量，分別為 X_1 與 X_2 。
- (2) 視這兩變量為羅吉斯模型(logistic model)的解釋變項，藉以生成真實的傾向分數PS。
- (3) 以此傾向分數作為指派組別(Z)的真實機率，來產生實驗組別。
- (4) 就這兩變量實驗組別的數據進行羅吉斯迴歸，取配適值為觀察的傾向分數 \widehat{PS} 。

亦即

$$\widehat{PS} = P(Z|X_1, X_2) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}$$

其中

$$\left\{ \begin{array}{l} \text{兩變量分配: } X_1 \sim N(0,1), X_2 \sim \text{Ber}(0.5) \\ \text{各係數設定}^2: \beta_0 = -2, \beta_1 = \beta_2 = \ln(2+t), t = 0,1, \dots, 8 \\ \text{處理指標: 實驗組 } Z = 1, \text{ 對照組 } Z = 0 \end{array} \right.$$



² 截距項係數 β_0 之正負影響實驗組與對照組各數的多寡，負數表示實驗組個數少於對照組，因估計重在實驗組平均處理效果(ATT)，故這邊皆設定-2。變量的係數(β_1, β_2)分配結構相近程度，由 $\ln 2$ 到 $\ln 10$ 兩組結構關係由高到低。

第四章 研究結果與討論

本章將就模擬實驗結果作分析，依序為臨界值求取、檢定力比較，最後是不同分配差異下的配對數前後比較。所有模擬實驗皆透過 R 軟體來進行。

第一節 臨界值的選定

本文所選用的模型為 $N(0,1)$ 、 $SN(0,1,-10)$ 、 $Fréchet(0,1,5)$ ，模擬結果如下表 3-2-1.1

所示：

表 4-1.1 95% 臨界值

Distributions	statistics percentile	KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
		95%	95%	95%	95%	95%	95%	95%
$N(0,1)$	n=1000	0.0146	0.0146	0.0144	0.1202	0.1198	0.1194	0.0600
	n=500	0.0247	0.0246	0.0245	0.1592	0.1598	0.1586	0.0840
	n=200	0.0506	0.0504	0.0495	0.2364	0.2374	0.2343	0.1350
$SN(0,1,-10)^*$	n=1000	0.0136	0.0136	0.0100	0.1093	0.1094	0.1092	0.0610
	n=500	0.0233	0.0231	0.0175	0.1445	0.1446	0.1443	0.0840
	n=200	0.0478	0.0494	0.0340	0.2053	0.2064	0.2049	0.1350
$Fréchet(0,1,5)^*$	n=1000	0.0182	0.0180	0.0139	0.1185	0.1177	0.1177	0.0600
	n=500	0.0307	0.0309	0.0248	0.1589	0.1589	0.1579	0.0860
	n=200	0.0596	0.0600	0.0437	0.2227	0.2239	0.2218	0.1350

*由於 KLIC 在文獻中提到欲比較之兩分配必須建構在相同定義域下才可積分，否則積分值會發散導致無效的 KLIC 值，故實務操作上，因模擬樣本之 PDF 受樣本大小、偏態不同而有不同的定義域，因而作兩點模擬積分處理：(1) 模擬樣本分配函數時，將超過左右兩端點的函數值延伸，認為可行的想法是因為越接近末端其機率近乎 0，且兩樣本來自相同母體其定義域不會差距太大，故延伸最末端值可使積分有效且影響不大；(2) 在左右偏分配下有少量 (< 2%) 的資料會特別延伸，有時使得此延伸端的離群值與大多資料中間機率值為 0，同樣會使積分無效，故分別對兩偏態較長一端取 99.9 百分位數進行積分，其認為可行的原因與第一點類似，認為左右偏下較長一端的機率值理論應非常小，而會產生過遠的分布使中間的機率斷層表示末端的機率非常小且分布不多，故適當的百分選取可以讓積分有效且不過分影響積分值。處理的方式應用在第二節的檢定力比較。

由上表可看出不論哪種分配各個 KLIC 與 KS 統計量的臨界值皆隨著樣本數遞增而變小，顯示出樣本數越少，各統計量的變異也越大。故樣本數較少的情況下，各統計量的數值需要相對的大，才能夠反映實驗與對照兩分配間的差異。

另外比較三種分配可以發現，在相同樣本數之下，分配不同對臨界值並沒有太大影響，但相較 KS 來說，KLIC 似乎較不穩定，我們認為其原因可能是因為 KS 僅著重最大累積機率誤差，其差異最多為 1，而 KLIC 討論的是整體的機率密度平均誤差，並沒有上限限制，故在左右偏態時可能因不同形狀參數造成偏峰態不同，而影響積分值大小。三個分配(常態、的偏態係數與峰態係數近似值如表 4-1.2 所示，可看出Fréchet明顯高出許多，所以針對不同偏峰態圖形所對應到的 KLIC 也有些許不同，但在大樣本下基本上是穩定的。

表 4-1.2 各分配之偏峰態係數比較

分配	$N(\mu, \sigma)$	$SN(\mu, \sigma, \alpha = -10)$	$Fréchet(\mu, \sigma, \alpha = 5)$
偏態	0	$\frac{4-\pi}{2} \frac{\left(\delta \sqrt{\frac{2}{\pi}}\right)^3}{\left(1-\frac{2\delta^2}{\pi}\right)^{3/2}} *$	$\begin{cases} \frac{\Gamma\left(1-\frac{3}{\alpha}\right)-3\Gamma\left(1-\frac{2}{\alpha}\right)\Gamma\left(1-\frac{1}{\alpha}\right)+2\Gamma^3\left(1-\frac{1}{\alpha}\right)}{\sqrt{\left(\Gamma\left(1-\frac{2}{\alpha}\right)-\Gamma^2\left(1-\frac{1}{\alpha}\right)\right)^3}} & \alpha > 3 \\ \infty & o.w \end{cases}$
峰態	0	$2(\pi-3) \frac{\left(\delta \sqrt{\frac{2}{\pi}}\right)^4}{\left(1-\frac{2\delta^2}{\pi}\right)^2} *$	$\begin{cases} -6 + \frac{\Gamma\left(1-\frac{4}{\alpha}\right)-4\Gamma\left(1-\frac{3}{\alpha}\right)\Gamma\left(1-\frac{1}{\alpha}\right)+3\Gamma^2\left(1-\frac{2}{\alpha}\right)}{\sqrt{\left(\Gamma\left(1-\frac{2}{\alpha}\right)-\Gamma^2\left(1-\frac{1}{\alpha}\right)\right)^3}} & \alpha > 4 \\ \infty & o.w \end{cases}$

將生成資料所預設的參數代入，得到 Skew-Normal 的偏態係數與峰態係數各為-9.5556與0.8232；

Fréchet偏態係數與峰態係數則為3.5350 和 12.6859；常態分配則都是 0。

* $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ 。

第二節 型一錯誤機率及檢定力

為看出檢定力之變化，我們分別變動位置、散佈及形狀參數的參數值。以表 4-2.1、4-2.2、及 4-2.3 呈現的是在樣本數為 1000 的情形下，對稱的常態分配、左偏的偏斜常態分配、及右偏的Fréchet分配在信心水準 5%之下，檢定力之變化。

表 4-2.1 樣本數 1000 時，常態分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0144	0.0144	0.0143	0.1192	0.1192	0.1187	0.0600
α^{***}	N(0,1)	0.0517	0.0554	0.0535	0.0543	0.0537	0.0546	0.0525
Location	N(0.1,1)	0.3359	0.3294	0.3435	0.3224	0.3196	0.3291	0.4777*
	N(0.2,1)	0.9313	0.9325	0.9379	0.9205	0.9192	0.9268	0.9722*
	N(0.3,1)	0.9998	0.9997	0.9998	0.9997	0.9997	0.9997	1.0000 *
	N(0.4,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	N(0.5,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	N(0.7,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	N(1,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Scale	N(0,1.1)	0.5666	0.6479*	0.6263	0.3515	0.5978	0.4939	0.1739
	N(0,1.2)	0.9966	0.9972*	0.9972	0.9475	0.9954	0.9859	0.6790
	N(0,1.3)	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000	0.9848
	N(0,1.4)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	N(0,1.5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	N(0,2)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	N(0,3)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	t(500,0)	0.0365	0.0291	0.0127	0.0366	0.0367	0.0371	0.0480*
	t(100,0)	0.0586*	0.0216	0.0150	0.0383	0.0412	0.0400	0.0478
	t(10,0)	0.8789*	0.0196	0.1447	0.0787	0.1384	0.1087	0.0588
	t(7,0)	0.9919*	0.0233**	0.3165	0.1473	0.2724	0.2152	0.0859
	t(4,0)	1.0000*	0.0150	0.7430	0.5404	0.7049	0.6539	0.2922
	t(2,0)	0.9919**	0.0144	0.9974	0.9991	0.9993	0.9995*	0.9991
	t(1,0)	0.1952**	0.8654**	1.0000	1.0000	1.0000	1.0000	1.0000

第四列開始每七列為一種參數變動，由上而下依序為位置、變異、厚尾的變化。

*最高檢定力、**異常值、***信心水準。

為確保信心水準 α 維持在 0.05，表中第三列呈現的是樣本同樣來自 $N(0,1)$ 的情況。由於 10000 次模擬實驗中，檢定統計量大於臨界值的比率，都十分靠近 0.05，與理論值吻合。再者，針對位置與散佈參數的變動，六種 KLIC 與 KS 統計量都有極高的檢定力，大體上只要參數的變化不要太小，檢定力幾乎可以達到 100%。Dowd(2005, p345-p347) 中提到 KS 對位置參數變動較敏感，對於尾端分配變動的反應效果較差，表中數據確實也反映出這種現象[11]。由表中可以看到同樣在位置參數變化下，KLIC 檢定力雖不比 KS 高，卻差異不大，但在散佈參數時 KLIC 表現比 KS 好許多。形狀參數(也可說尾部薄厚程度)的部分，各個統計樣的表現較沒有一致性，檢定的效果也較不理想，其中紅色數據³為不符合預期的模擬結果，全部出現在 KLIC 的情況。這顯示出 KLIC 在厚尾的狀態下其檢定力並不穩定，究其原因，我們猜測可能是因為機率的互消或是積分範圍不夠完整產生負值造成檢定力降低，而第三種 KLIC 是修正 KLIC1 和 KLIC2 所得，比較不會有過度重視某一方而偏頗的情形，以及 KLIC4 到 KLIC6 的絕對值修正皆較單純 KLIC1 和 KLIC2 來的高檢定力。

³其積分實務操作上修正方式與第壹部分相同，惟厚尾區段 t 分配選取以 0.5% 和 99.5% 百分位數作為積分上下界，利用上一節原理，少數時候會有過遠的資料使得在離群與大多數資料之間形成近乎 0 的機率斷層使積分發散，而多數這種現象發生都幾乎有一個共通性—斷層的末端樣本個數比非常少 (< 2%)，故以相同方式處理。

表 4-2.2 樣本數 1000 時，(左)偏斜常態分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0136	0.0136	0.0100	0.1093	0.1094	0.1092	0.0610
α^{**}	SN(0,1,-10)	0.0476	0.0470	0.0512	0.0484	0.0488	0.0474	0.0446
Location	SN(-0.01,1,-10)	0.0428	0.0653*	0.0587	0.0547	0.0549	0.0558	0.0547
	SN(-0.02,1,-10)	0.0447	0.0987*	0.0857	0.0810	0.0801	0.0798	0.0874
	SN(-0.03,1,-10)	0.0626	0.1634*	0.1570	0.1344	0.1260	0.1298	0.1556
	SN(-0.05,1,-10)	0.1700	0.3972	0.4457*	0.3447	0.3017	0.3259	0.4062
	SN(-0.08,1,-10)	0.5848	0.8056	0.9125*	0.7950	0.6935	0.7527	0.8531
	SN(-0.1,1,-10)	0.8672	0.9457	0.9914*	0.9604	0.8985	0.9367	0.9778
	SN(-0.2,1,-10)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Scale	SN(0,1.05,-10)	0.2921	0.0392	0.1755*	0.1268	0.1661	0.1460	0.1745
	SN(0,1.1,-10)	0.7689*	0.1085	0.5578	0.4264	0.5250	0.4795	0.5607
	SN(0,1.15,-10)	0.9803*	0.2431	0.8978	0.8084	0.8778	0.8519	0.8984
	SN(0,1.2,-10)	0.9997*	0.3891	0.9882	0.9689	0.9833	0.9789	0.9907
	SN(0,1.3,-10)	1.0000	0.6167	1.0000	0.9999	1.0000	1.0000	1.0000
	SN(0,1.4,-10)	1.0000	0.7306	1.0000	1.0000	1.0000	1.0000	1.0000
	SN(0,1.5,-10)	1.0000	0.7805	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	St(0,1,-10,500)	0.0936*	0.0781	0.0399	0.0362	0.0356	0.0356	0.0444
	St(0,1,-10,100)	0.1356*	0.0598	0.0439	0.0382	0.0384	0.0383	0.0428
	St(0,1,-10,10)	0.9200*	0.0046	0.2490	0.1037	0.1496	0.1243	0.1863
	St(0,1,-10,7)	0.9943*	0.0032	0.4344	0.2082	0.2888	0.2484	0.4179
	St(0,1,-10,5)	1.0000*	0.0018	0.6790	0.4475	0.5363	0.4966	0.7967
	St(0,1,-10,4)	1.0000*	0.0007	0.8332	0.7212	0.7537	0.7468	0.9654
	St(0,1,-10,3)	1.0000*	0.0004	0.9610	0.9628	0.9545	0.9613	0.9999

為使積分有效，在左端同樣選取0.5%百分位數作為積分下界。

*最高檢定力、**信心水準

表 4-2.2 的結果大致上與對稱分配情況相似。針對位置參數變化，各個統計量都有較高的檢定力；其次是散佈參數，最差是形狀參數。整體而言，最高檢定力出現以 KLIC1 居多，除 KLIC2 在形狀參數是呈現遞減的趨勢之外，其他皆是隨著分配差異增加而上升。第四到第六類經由絕對值調整過後的 KLIC 統計量，雖然檢定力略小於前三類中之最大者，但三者間的差異不大，顯示穩定性較高。另外 KS 雖不為最高檢定力，但落後程度不大，數值上與後三類的 KLIC 差不多。同樣的，表中第三列呈現的仍是為錯誤拒絕 H_0 的機率，亦即信心水準。

表 4-2.3 樣本數 1000 時，Fréchet 分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0182	0.0180	0.0139	0.1185	0.1177	0.1177	0.0600
α^{**}	Fréchet(0,1,5)	0.0506	0.0520	0.0519	0.0474	0.0514	0.0489	0.0528
Location	Fréchet(0.02,1,5)	0.1376	0.2478	0.2950	0.3174	0.3040	0.3138	0.3919*
	Fréchet(0.03,1,5)	0.3667	0.5195	0.6707	0.6631	0.6179	0.6475	0.7532*
	Fréchet(0.05,1,5)	0.9263	0.9497	0.9943	0.9915	0.9778	0.9873	0.9966*
	Fréchet(0.08,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0.1,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0.15,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0.2,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Scale	Fréchet(0,1.02,5)	0.1521	0.1792	0.2580	0.2748	0.2973	0.2886	0.3999*
	Fréchet(0,1.05,5)	0.8855	0.8348	0.9769	0.9737	0.9654	0.9711	0.9935*
	Fréchet(0,1.1,5)	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.15,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.2,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.25,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.3,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	Fréchet(0,1,4.8)	0.1771*	0.0328	0.0964	0.0796	0.1206	0.0998	0.0648
	Fréchet(0,1,4.5)	0.6932*	0.0758	0.4574	0.3770	0.5273	0.4616	0.2131
	Fréchet(0,1,4)	0.9998*	0.4937	0.9935	0.9870	0.9961	0.9942	0.8669
	Fréchet(0,1,3.5)	0.9999	0.9452	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1,3)	0.9999	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1,2.5)	0.9988**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1,2)	0.9976**	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

同樣為使積分有效，在右端選取99.5%百分位數作為積分下界。

**此處的是數值異常是因為Fréchet有太多的離群值位於右方遠端，造成積分範圍的不完整而產生負值，基本上差異仍是非常小。

*最高檢定力、**信心水準。

由表 4-2.3 可以發現，整體而言，所有統計量的檢定力都會隨著參數值的變動呈現出規則性的變化：參數差異越大，檢定力越高。不過依舊是在位置與散佈參數變動的情況表現較佳，在形狀參數實則遜色一些。值得一提的是，與前述左偏分配的變動情形作比較，右偏分配情形的檢定力效果除了具有規則性外，數值似乎也比左偏的情況來的大，尤其是 KS 統計量的表現相對耀眼許多。其原因可能出現在Fréchet本身分配受參數變化

影響較 Skew-normal 來的多，且兩者分配並沒有左右對照的關係，相較之下Fréchet明顯右偏許多。如圖 4-2 所示，對同樣移動 0.5 單位的形狀參數與散佈參數來看，Skew-normal 的 CDF 最大差異不似Fréchet大；但在形狀參數上變化就平分秋色，所以檢定力也反映出這樣的現象。同樣的，表中第三列用來檢驗信心水準，結果與理論值0.05十分接近。

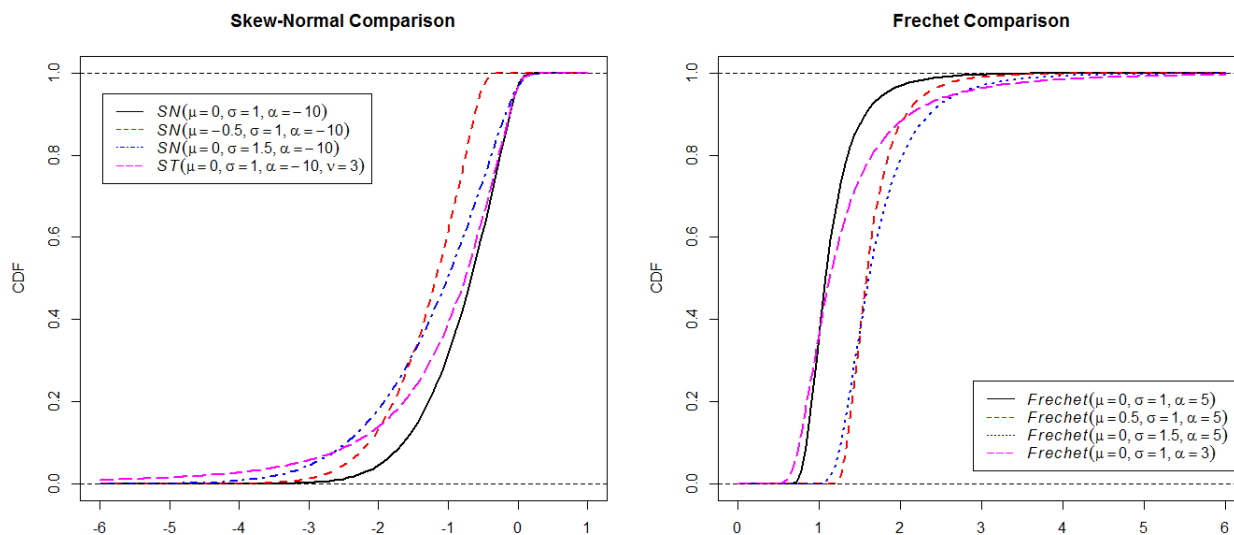


圖 4-2 左右偏態之 CDF 比較

我們想知道兩統計量在小樣本下的表現狀況及其穩定程度，故重複一樣實驗，將樣本數改為 500 與 200 去看檢定力變化，並將結果置於附錄表 1 至表 6，其討論內容如下。

整體而言，不論基準分配為何，一如預期檢定力隨樣本數增加而上升，基本上只要參數的差異到達某個程度，KLIC 或 KS 統計量都有辦法分辨出差異。進一步分析可以發現，對稱分配的情況下，三種樣本數都呈現出 KS 在位置參數的變化上有較強的檢定力。至於散佈參數與形狀參數的變動上，最高檢定力都出現在前三類 KLIC，但缺點是在形狀參數的檢定力較不具有規則性(尤其是 KLIC2)，推測其原因可能與 KLIC 在實務操作上的困難(積分範圍定義與選取)有關。相較之下，後三類 KLIC 的表現雖略遜於 KS 與前三類 KLIC，但是並不至於差太多，優點是規則性較強。其次，藉由左偏的偏斜常態分配可以發現，大致上前三類 KLIC 在各參數變動的檢定力表現都優於其他兩者，但

是與對稱分配相同，在形狀參數上其檢定力較不具規則性。比較出乎意料的是，後三類 KLIC 及 KS 在小樣本($n = 200$)時，針對形狀參數變動的檢定力不甚理想，在前三類 KLIC 可以達到 98% 的檢定力時，KS 只有 40%，而後三類 KLIC 更僅 27%。至於右偏 Fréchet 分配的情況，這些統計量的表現最合乎預期，我們猜測這或許與分配參數變動時較其他兩者來的明顯有關。如同前面提到，KS 對形狀參數變動較不敏感，因此其檢定力較六種 KLIC 來的低。此外，由於此處的形狀參數為分配的既有參數，不像是常態與(左)偏斜常態的情況，以 t 分配來描述厚尾分配的情形，故 KLIC 在積分上較沒有定義範圍的問題，較不易發生計算上的問題。

綜合以上九個表格，我們認為若分配具有一定差異，基本上 KLIC 與 KS 都可以準確判斷。但若要偵測一些較輕微的變化，針對不同參數的選擇會有所不同。雖然 KS 對於位置參數較為敏感，但若要判斷散佈參數與形狀參數時，KLIC 會有較佳的檢定力。不過在使用 KLIC 時，有個前提—欲比較的兩者分配在定義與(分布範圍)上要盡量一致，或者不能相差太遠，尤其我們是在比較兩實際樣本的分配時會受限更多因素。幸運的是，以傾向分數在配對上的實際應用，其分配在定義域都有一定重疊的比例，對於積分範圍選取較無前述問題。故綜合各種考量，我們認為可以考慮使用 KLIC3、KLIC6 以及 KS 統計量來作為平衡標準，主要原因是在於雖然不比 KLIC1 在許多狀況下有較高的檢定力，但此三者有較良好的規律性可以遵循，來減少診斷時產生的確定性問題。下一節將針對此三種統計量進行配對數比討論。

第三節 配對前後對照表

接著我們想知道分配結構差異所反映的 KLIC 與 KS 統計量與配對樣本比的可能關聯。在進行實際配對之前，我們先設定迴歸係數由 $\ln 2$ 變動到 $\ln 10$ 藉以產生不同分配型態之樣本共 10000 筆，計算其配對前後 KLIC1 到 KLIC6、KS 統計量以及配對比率，並做其散佈圖觀察趨勢，如下圖 4-3 所示。

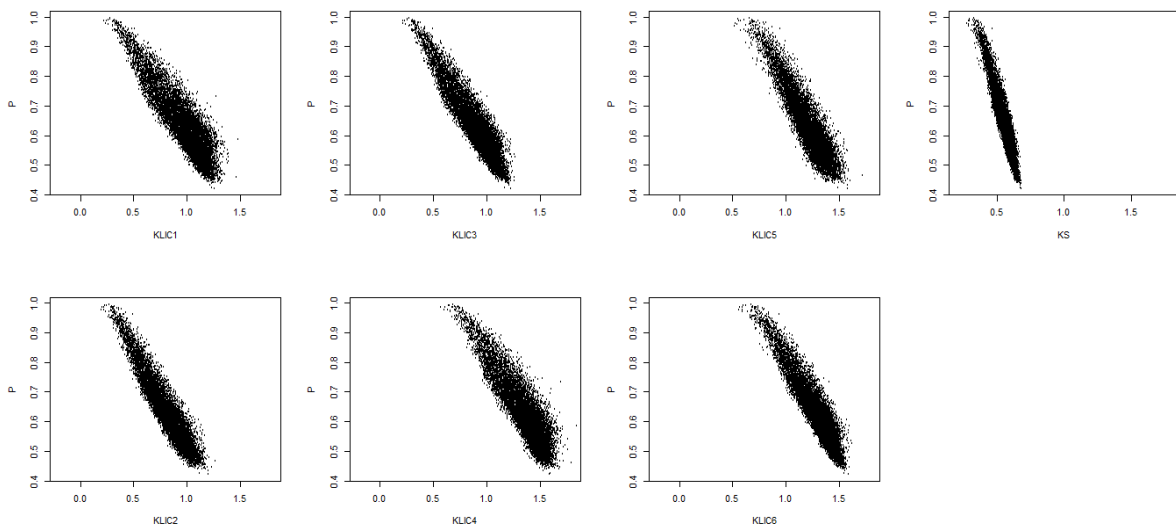


圖 4-3 各統計量與配對比率散佈圖

圖 4-3 由上排左邊數來為分別為 KLIC1、KLIC3、KLIC5、KS，下排則是 KLIC2、KLIC4、KLIC6 等統計量對配對比例的散佈圖，可觀察到配對數比例與此七個統計量皆呈現負相關的現象。大體上而言，KLIC 或 KS 值越大，相對的配對數比例會變少。雖然此七個散佈圖看似十分接近，但 KLIC2、KLIC3、KLIC6 以及 KS 似乎具有較高的相關程度。相對應的皮爾森相關係數及各個統計量的全距，如表 4-3.1 所示。其中相關係數皆近似 -0.9 ，KS 最高，其值為 -0.94 ，這顯示出這幾個統計量皆和配對樣本比呈現高度負相關。考量檢定力的相關訊息後，我們選擇以 KLIC3、KLIC6、KS 去做進一步的比較說明⁴，其他統計量比較請參照附錄表 7 到表 9。

⁴ 雖然 KLIC 在相關係數上有較強的相關，但是由於其變動不穩定，故僅將其變動置於附錄不討論。

表 4-3.1 各統計量的全距與配對比率之相關係數

Statistics	Correlation coefficient (vs. P)	Range
KLIC1	-0.9031	1.2563
KLIC2	-0.9366	1.0652
KLIC3	-0.9294	1.0569
KLIC4	-0.8941	1.2689
KLIC5	-0.9109	1.1956
KLIC6	-0.9296	1.0558
KS	-0.9408	0.4650

表 4-3.2 中，第一排為各變量之係數(β_1, β_2)與配對前後的列表，第二到四排為三統計量(KLIC3、KLIC6、KS)在不同樣本數(2000、1000、500)情況下的平均數，最後一欄為三種樣本數下的平均配對比。比照表中數值可以發現，整體而言，隨著變量係數的增加，表示原始兩組分配差異越大，此時各統計量皆呈現上升的趨勢。再對照最右排配對樣本前後比，可看到配對比跟著下降，顯示可能配對成功的樣本數受一開始分配結構所影響。接著我們去細分各個統計量與配對比例，先將配對比例依十分位數劃分，如表中紅色實線所示，可發現即使相同統計量，卻仍受樣本數配對比例的影響而有輕微的不同，不過大抵上仍可以區分出約略的範圍作為判斷依據。例如我們可把配對比例低於 50% 做為一個參考指標，當 KLIC3 高於 0.93 時，我們可以預期原始樣本中有極大的比例可能無法進行配對。使用 KS 也能達到相同目的，可是由於 KS 本身有上界限制，其值域範圍較窄，其差異在初始變化大，而後又太過微小，對於作為比例對照可能判斷困難，故此處我們建議選用 KLIC3 或 KLIC6 做為一個配對比的指標。

表 4-3.2 各樣本數之配對前後各統計量平均數比較與配對比例

Test Statistics		KLIC3			KLIC6			KS			Proportion of matched No.		
Sample No.		2000	1000	500	2000	1000	500	2000	1000	500	2000	1000	500
coefficient	match	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ	μ
ln 2	Before	0.2815	0.2695	0.2782	0.6561	0.6295	0.6349	0.3074	0.3178	0.3328	0.9831	0.9773	0.9678
	After	0.0005	0.0007	0.0009	0.0082	0.0105	0.0146	0.0142	0.0221	0.0358			
ln 3	Before	0.5922	0.5347	0.5310	1.0004	0.9362	0.9282	0.4309	0.4386	0.4500	0.8720	0.8680	0.8594
	After	0.0006	0.0010	0.0012	0.0202	0.0219	0.0254	0.0258	0.0339	0.0465			
ln 4	Before	0.8021	0.7083	0.6930	1.1994	1.1083	1.0894	0.4989	0.5056	0.5151	0.7531	0.7503	0.7454
	After	0.0009	0.0014	0.0017	0.0311	0.0322	0.0353	0.0344	0.0431	0.0560			
ln 5	Before	0.9365	0.8240	0.8023	1.3210	1.2153	1.1904	0.5429	0.5485	0.5577	0.6659	0.6641	0.6593
	After	0.0012	0.0019	0.0022	0.0402	0.0409	0.0436	0.0409	0.0501	0.0636			
ln 6	Before	1.0271	0.9080	0.8787	1.3989	1.2894	1.2581	0.5740	0.5794	0.5875	0.6029	0.6013	0.5983
	After	0.0015	0.0025	0.0028	0.0482	0.0486	0.0506	0.0462	0.0559	0.0693			
ln 7	Before	1.0905	0.9710	0.9411	1.4519	1.3429	1.3112	0.5972	0.6022	0.6114	0.5560	0.5542	0.5501
	After	0.0019	0.0030	0.0034	0.0553	0.0555	0.0575	0.0509	0.0609	0.0752			
ln 8	Before	1.1378	1.0211	0.9870	1.4901	1.3851	1.3488	0.6156	0.6208	0.6282	0.5194	0.5180	0.5144
	After	0.0021	0.0036	0.0039	0.0613	0.0612	0.0634	0.0548	0.0648	0.0799			
ln 9	Before	1.1750	1.0638	1.0289	1.5195	1.4194	1.3832	0.6308	0.6359	0.6434	0.4900	0.4884	0.4857
	After	0.0025	0.0041	0.0045	0.0671	0.0670	0.0685	0.0586	0.0689	0.0836			
ln 10	Before	1.2055	1.0993	1.0638	1.5437	1.4484	1.4112	0.6436	0.6483	0.6560	0.4663	0.4651	0.4618
	After	0.0028	0.0047	0.0050	0.0722	0.0719	0.0732	0.0617	0.0725	0.0878			

表 4-3.3 不同樣本之配對與 KLIC 計算時間比較

Test Statistics		N=2000		N=1000		N=500	
Time-consuming		KLIC*4	PSM	KLIC*4	PSM	KLIC*4	PSM
ln 2	μ	0.0070	0.0290	0.0067	0.0083	0.0067	0.0030
	σ	0.0002	0.0001	0.0002	0.0002	0.0002	0.0001
ln 3	μ	0.0064	0.0360	0.0064	0.0096	0.0063	0.0033
	σ	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001
ln 4	μ	0.0059	0.0403	0.0058	0.0107	0.0058	0.0035
	σ	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001
ln 5	μ	0.0057	0.0428	0.0055	0.0112	0.0052	0.0037
	σ	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001
ln 6	μ	0.0056	0.0447	0.0051	0.0116	0.0051	0.0038
	σ	0.0002	0.0001	0.0002	0.0002	0.0002	0.0001
ln 7	μ	0.0053	0.0462	0.0050	0.0119	0.0049	0.0039
	σ	0.0002	0.0001	0.0002	0.0002	0.0002	0.0001
ln 8	μ	0.0051	0.0473	0.0049	0.0122	0.0048	0.0038
	σ	0.0002	0.0001	0.0002	0.0002	0.0002	0.0001
ln 9	μ	0.0049	0.0482	0.0048	0.0124	0.0046	0.0039
	σ	0.0002	0.0001	0.0002	0.0002	0.0001	0.0001
ln 10	μ	0.0048	0.0489	0.0046	0.0125	0.0046	0.0039
	σ	0.0002	0.0001	0.0002	0.0002	0.0002	0.0001

中間三排為不同樣本數的時間計算，網底部分為平均數，KLIC*4 表示計算一次 KLIC1、KLIC2、KLIC4、KLIC5 所需的時間，單位為秒，PSM 表示配對一次所需的時間。舉例來說，第三行第五列 0.0058 表示在兩變量係數皆為 ln 4 之下，計算總樣本數 1000 筆配對前 KLIC1、KLIC2、KLIC4、KLIC5 平均而言需要 0.0058 秒，計算 KLIC3 與 KLIC6 則需要 0.0029 秒。

最後我們將比較傾向分數配對與計算 KLIC 所花費的時間，用以說明 KLIC 比配對更具有經濟效應。表 4-3.3 為前述模擬時過程所花費的時間，我們可以看到左邊第一排為模擬所預設的各個比較係數與對照的平均數和標準差，而中間三排為樣本數 2000、1000、500 的四種 KLIC 計算時間與配對時間，觀察發現在樣本數少的時候配對時間並沒有特別明顯，但隨著樣本增加 2 倍，其所需耗費的時間曾為原來的三倍，可由表 4-3.3 右到左看出來，相較之下 KLIC 計算並不隨著樣本數增加而有明顯的波動，故認為在樣本數過大時，如巨量資料，KLIC 或許可以當作檢驗是否必須去做傾向分數配對的一個先行指標。

第五章 實證分析

第一節 主題分析

本章主要探討「帕金森氏症與跨步時間的關係」。

帕金森氏症(Parkinson's disease (PD))是最常見的神經病變之一，屬於進行性動作障礙疾病的一種，在美國每年約有 6 萬人口被診斷罹患此症狀，其中 60 歲以上老年人較為常見。其發病具有階段性，常見症狀為僵直、震顫、運動遲緩、平衡及行走障礙等等，根據嚴重程度依 Hoehn&Yahr 分級表分為五個階段，也可能伴隨其他症狀如抑鬱、記憶力減退、語言障礙等等，目前主要以藥物治療和手術治療為主。

帕金森氏症並沒有特異性診斷試驗，在臨床上將帕金森氏症分為三類，分別是原發性、非典型、次發性。研究上多以跨步時間(stride time)、擺盪時間(swing time)等步態表現搭配跑步機與感測器來測量，Frenkel-Toledo, Giladi, Peretz, et al.(2005)文中提到帕金森氏症病患在正常速度下步態表現的變異較正常病患大，並指出步態表現變異與罹患帕金森氏症有關[12]。

本章使用 Physionet 所提供研究帕金森氏症者其步態紀錄資料，利用傾向分數對兩組試驗者進行分配，並觀察其平衡效果[13]。

第二節 資料來源與變數定義

壹、資料來源

擷取 Physionet 中帕金森氏症步態研究所提供之提資料庫，正常速度下感測器測量的步態紀錄，包含 93 位 Hoehn&Yahr 分級 2 至 3 的帕金森氏症病患，以及 72 位無罹患帕金森氏症者，其中帕金森氏病患有效樣本 87 份，健康者有 71 份。我們想就此資料，

藉由傾向分數配對方法來進行實證分析。分析過程中，主要著眼於藉由配對前之 KLIC 及 KS 值來預估配對比率，並與真實配對比率做比較；以及配對後 KLIC 及 KS 值的計算，作為平衡診斷的一個判斷準則。

貳、變數說明

以有無罹患帕金森氏症作為試驗分組的指標(Z_i)進行配對，詳述如下：

$$Z_i = \begin{cases} 1, & \text{罹患帕金森氏症(暴露組)} \\ 0, & \text{未患帕金森氏症(非暴露組)} \end{cases}$$

並以調整過的跨步時間變異(stride time variability(STV), Y_i)為評分標準，所包含變數共七個，分別為性別、年齡、身高、體重、步行速度、平均步行時間、跨步時間標準差。

後兩項變數為感測器所測得之數據整理，受測者在水平地面以各自正常速度行走兩分鐘，並記錄雙腳各 8 個感測器所呈現之地面垂直反作用力(vertical ground reaction force, VGRF)，來計算腳尖離地與著地之時間以換算步行時間，紀錄頻率為 100 赫茲。

第三節 分析結果

壹、分析步驟

實證分析步驟如下：

- (1) 傾向分數估計：我們依上述變量進行羅吉斯迴歸(logistic regression)以估計傾向分數。
- (2) 配對前比較：繪製配對前傾向分數分配之對照圖，並分別計算配對前六種 KLIC 與 KS 統計量，用以推測成功配對的比率。雖然第四章中我們建議採用 KLIC3 或 KLIC6 作為平衡診斷工具，不過為了進行比較，我們仍將七個統計量的結果全數列出。

- (3) 配對方式：以傾向分數作為配對依據，採用貪婪分配，並設定門檻(caliper)值為 0.2。
- (4) 配後比較：計算成功配對的比率，繪製配對後傾向分數分配之對照圖，並計算配對後六種 KLIC 與 KS 統計量之數值，進行平衡診斷。

貳、分析結果

我們將配對前後的 KLIC 與 KS 值整理於表 5-3.1。就配對前各統計量的數值與表 4-1.1 的臨界值做比較後可以發現，除 KS 略大之外，其他六個 KLIC 皆較 95% 臨界值低，顯示其原始資料分布應差異不大，對照圖 5-3.1 左圖可知分配大致相似，但仍有些許落差，可能是樣本數太小所致。雖然表 4-1.1 的臨界值是架構於三類特定的機率分配，此資料傾向分數呈現雙峰分配未必適用，但我們仍進一步分析配對前後之差異。圖 5-3.1 中呈現的是配對前後兩組之傾向分數分配的對照圖，其中黑色實線為暴露組，綠色虛線則是非暴露組。

由於原始資料樣本數總數 158 筆，我們以表 4-3.2 中最接近的情況依樣本數 500 做為參考，推估成功配對比例應高於 0.97，表示配對成功比率非常高，應該值得進行傾向分數配對。由表 5-3.2 可以發現實際進行配對後，71 個非暴露組的試驗對象中，69 人成功配對，配對比例接近 0.97，某種程度上符合預期。

最後，由表 5-3.1 可以發現，對照配對前與配對後各統計量皆明顯下降許多，也遠低於表 4-1.1 中之臨界值，這意味著配對後暴露與非暴露兩組已可視為達到平衡。圖 5-3.1 的右圖也反映此項事實，該圖呈現的是配對後暴露與非暴露兩組之傾向分數分配的對照圖。由於兩組之傾向分數機率密度函數圖幾乎一模一樣，顯示出傾向分數配對之後兩者的背景變項的分配幾乎相同，因此可視為隨機配置。另外，由表 5-3.1 中也可以看到配對後的 KLIC1 出現非常小的負號，與模擬結果類似，是在計算 KLIC 過程中若兩樣本非常接近時，因其定義在實務上的處理使得負號產生，在此可視為兩者分配非常接近。

表 5-3.1 配對前後統計量

	KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
配對前	0.0162	0.0162	0.0162	0.1245	0.1246	0.1246	0.1470
配對後	-0.0000*	0.0000	0.0000	0.0000	0.0000	0.0000	0.0290

*實務操作上太相近的兩個分配其 KLIC 值計算上受範圍選取而產生負值。
上述配對後 KLIC2 至 KLIC6 皆為小於 0.00001 之正數。

表 5-3.2 配對前後非暴露組個數與比例

N=158	配對前	配對後	比例
非暴露組	71	69	0.9718

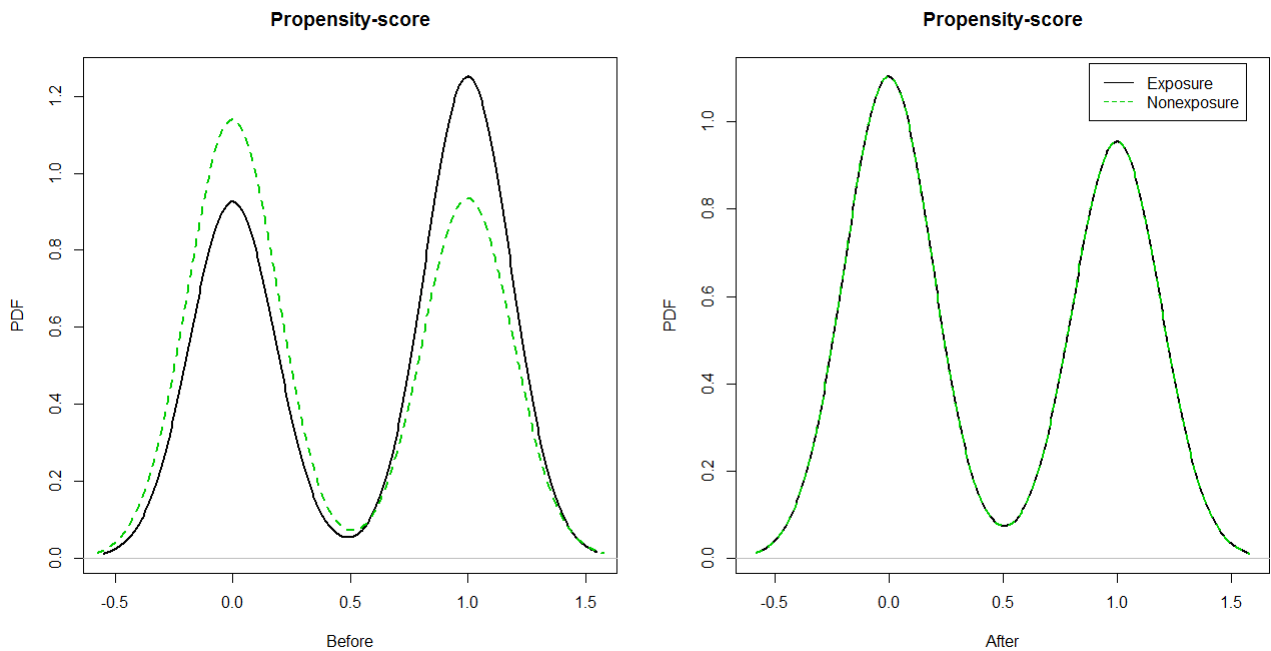


圖 5-3.1 配對前後傾向分數 PDF 比較

第六章 結論與建議

第一節 結論

傾向分數的概念已經被廣泛地應用在觀察性研究上，藉由傾向分數進行配對後，可以使兩組資料達到某種程度的平衡，獲致近似隨機化的配置效果，進而進行有效的統計推論。除了文獻中既有的平衡診斷工具主要聚焦於兩組資料間平均數及變異數的比較，本文中我們提出使用 KLIC 及 KS 兩種可以比較分配差異的統計量，作為平衡診斷的另一類工具。

KLIC 理論值為一正數，但在實際資料處理上，由於我們並不會知道真實分配為何，故就樣本來推斷真實分配除了本身具有誤差，在定義域上面選取也有所設限，造成 KLIC 計算上會有負值的出現，導致可能的偏誤現象。因此除了傳統的 KLIC1 統計量外，我們也考慮了五種變形，分別標記為 KLIC2、KLIC3、...、KLIC6。KLIC1、KLIC2 在兩者分配非常接近時，有比較高的機率會產生負號，但 KLIC3~KLIC6 則無此情況發生。在考量計算結果的穩定度，及型一、二錯誤的發生機率等因素後，我們建議可以使用 KLIC3 或 KLIC6 及 KS 作為平衡診斷的一種工具。KLIC 與 KS 其實可以一併使用，兩者一為比較分配的平均差異，一為比較分配的最大差異，在厚尾情況與特定少數跳點情況表現各有優缺，因此可以彌補彼此的不足。

由於後續統計推論的正確與否，取決於配對結果是否達到平衡。此外配對成功數的多寡，也關係到統計推論的適用範圍；配對成功比例過低，推論的適用範圍勢必縮小。然而配對的進行可能需要耗費一些時間，因此如果有個統計量可以做為配對前的先行指標，提供配對成功比例的一個估計，如此一來只要預測的成功配對數的比例過低，或許研究者就不需花費時間進行無謂的配對。模擬結果顯示，KLIC 或 KS 與配對成功比例間都呈現出負相關的現象，也就是說 KLIC 或 KS 的數值越大，配對成功比例明顯下降。由於 KS 的值域局限於 0、1 之間，範圍較窄，我們建議研究者在配對之先，或許可以透過 KLIC3 或 KLIC6 的計算先行判斷可能的配對比例，再來決定是否進行後續的配對。

另外原先考量到 KLIC 原始定義在操作上可能有正負相消的問題，故將此做一個絕對值的調整以預期表現能勝於原始 KLIC，但結果顯示雖然 KLIC1 和 KLIC2 的確有不穩定的情況，但是若以兩者平均的 KLIC3(也就是以 Jeffreys information 角度)來看，其表現卻兼具一致性與穩定性，況且，KLIC 原先具有良好的性質在加入絕對值修正後不一定存在，故若要選取還是建議以 KLIC3 優先。

第二節 未來研究方向建議

本文僅著重在兩變量與樣本數 2000、1000、500 的樣本上去探討配對比例與統計量之關係，雖然我們觀察到某種規律性，但這樣的結果仍不夠充分，往後或許可以進行更大規模的探討，或許可以獲致較為具體的發現。

此外，有關檢定力大小的探討，在模擬過程中，我們是以電腦亂數選取重新抽樣進行比較，然而實務處理時事實上只有一組數據，因此想要透過 KLIC 或 KS 來進行平衡與否的檢定時，或許可以透過排列檢定(permutation test)的方式來進行。

參考文獻

1. Rosenbaum, P.R. and D.B. Rubin, *The central role of the propensity score in observational studies for causal effects*. Biometrika, 1983. **70**(1): p. 41-55.
2. Austin, P.C., *An introduction to propensity score methods for reducing the effects of confounding in observational studies*. Multivariate Behavioral Research, 2011. **46**(3): p. 399-424.
3. Frölich, M., *Finite-sample properties of propensity-score matching and weighting estimators*. Review of Economics and Statistics, 2004. **86**(1): p. 77-90.
4. Busso, M., J. DiNardo, and J. McCrary, *New evidence on the finite sample properties of propensity score reweighting and matching estimators*. Review of Economics and Statistics, 2011(0).
5. Cover, T.M. and J.A. Thomas, *Entropy, relative entropy and mutual information*. Elements of Information Theory, 1991: p. 12-49.
6. Ullah, A., *Entropy, divergence and distance measures with econometric applications*. Journal of Statistical Planning and Inference, 1996. **49**(1): p. 137-162.
7. Kullback, S. and R.A. Leibler, *On information and sufficiency*. The Annals of Mathematical Statistics, 1951: p. 79-86.
8. Austin, P.C., *Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples*. Statistics in medicine, 2009. **28**(25): p. 3083-3107.
9. Azzalini, A., *The skew-normal and related families*. Vol. 3. 2013: Cambridge University Press.
10. Austin, P.C., *The performance of different propensity score methods for estimating marginal odds ratios*. Statistics in medicine, 2007. **26**(16): p. 3078-3094.
11. Dowd, K., *Measuring market risk*. 2007: John Wiley & Sons.
12. Frenkel-Toledo, S., et al., *Journal of NeuroEngineering and Rehabilitation*. Journal of neuroengineering and rehabilitation, 2005. **2**(23): p. 0003-2.
13. Goldberger, A.L., et al., *PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals*. Circulation, 2000. **101**(23): p. E215-20.

附錄

程式碼

```
library(Matching)
#傾向分數估計
PS<-glm(outcome ~ x1+x2+x3+x4+x5+x6,data=DATA,family=binomial(link=logit))$fitted
#傾向分數配對
PSM<-Match(Y=outcome, Tr=group, X=log(PS/(1-PS)), estimand="ATC", M=1,
ties=FALSE, replace=FALSE, caliper=0.2)

ks<-ks.test(group1,group2)$stat
#實際資料 pdf 估計
pi<-Vectorize(approxfun(density(groupi),rule=2)) #i=1,2
#KLIC1-6 函數
K1<-function(w){
  p1(w)*log(p1(w)/p2(w))}
K2<-function(w){
  p2(w)*log(p2(w)/p1(w))}
K4<-function(w){
  p1(w)*abs(log(p1(w)/p2(w)))}
K5<-function(w){
  p2(w)*abs(log(p1(w)/p2(w)))}
KLi<-integrate(Ki,min(c(group1,group2)),max(c(group1,group2)),rel.tol=Inf)$value
#i=1,2,4,5
KL3=(KL1+KL2)/2
KL6=(KL4+KL5)/2
```


表 1 樣本數 500 時，常態分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS	
95% Critical Value		0.0250	0.0252	0.0248	0.1620	0.1622	0.1615	0.0860	
α^{***}	N(0,1)	0.0489	0.0494	0.0496	0.0477	0.0475	0.0458	0.0429	
$\beta - \tau$	Location	N(0.1,1)	0.1792	0.1703	0.1812	0.1701	0.1675	0.1715	0.2441*
		N(0.2,1)	0.6459	0.6384	0.6576	0.6075	0.6070	0.6151	0.7480*
		N(0.3,1)	0.9659	0.9673	0.9687	0.9515	0.9516	0.9562	0.9804*
		N(0.4,1)	0.9997	0.9997	0.9996	0.9986	0.9991	0.9990	1.0000*
		N(0.5,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		N(0.7,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		N(1,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Scale	N(0,1.1)	0.2897	0.3557*	0.3351	0.1549	0.3313	0.2450	0.0856
		N(0,1.2)	0.8806	0.9025	0.9041*	0.6261	0.8900	0.8022	0.2805
		N(0,1.3)	0.9979	0.9982*	0.9982*	0.9523	0.9977	0.9922	0.6557
		N(0,1.4)	1.0000	1.0000	1.0000	0.9985	1.0000	0.9999	0.9268
		N(0,1.5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9942
		N(0,2)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		N(0,3)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Shape	t(500,0)	0.0344	0.0301	0.0149	0.0336	0.0325	0.0329	0.0417*
		t(100,0)	0.0464*	0.0209	0.0128	0.0327	0.0331	0.0326	0.0398
		t(10,0)	0.5492*	0.0167**	0.0643	0.0475	0.0774	0.0608	0.0500
		t(7,0)	0.8375*	0.0176	0.1237	0.0690	0.1233	0.0936	0.0504
		t(4,0)	0.9994*	0.0171	0.3283	0.1901	0.3017	0.2507	0.0967
		t(2,0)	0.9975*	0.0085	0.8120	0.8918	0.8436	0.8780	0.6759
		t(1,0)	0.5296**	0.5421**	0.9993	0.9968	0.9995	0.9991	1.0000*

*最高檢定力、**異常值、***信心水準

此表為第四章第二節，在樣本數 500 之下模擬 10000 次的常態分配 KLIC 與 KS 統計量之型一誤差率及檢定力。大致具有一致性，除 KLIC2 相反之外，其他六統計量皆隨著分配差異越大而上升，僅在厚尾分配表現上較差且不穩。

表 2 樣本數 500 時，(左)偏斜常態分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0233	0.0231	0.0175	0.1445	0.1446	0.1443	0.0840
α^{**}	SN(0,1,-10)	0.0457	0.0510	0.0496	0.0532	0.0520	0.0517	0.0524
Location	SN(-0.01,1,-10)	0.0398	0.0671*	0.0531	0.0545	0.0570	0.0555	0.0560
	SN(-0.02,1,-10)	0.0389	0.0877*	0.0671	0.0621	0.0662	0.0641	0.0711
	SN(-0.03,1,-10)	0.0391	0.1256*	0.0968	0.0895	0.0927	0.0908	0.1118
	SN(-0.05,1,-10)	0.0667	0.2474*	0.2048	0.1777	0.1756	0.1783	0.2325
	SN(-0.08,1,-10)	0.1976	0.5307	0.5301	0.4422	0.3816	0.4168	0.5406*
	SN(-0.1,1,-10)	0.3922	0.7308	0.7720*	0.6662	0.5694	0.6258	0.7641
	SN(-0.2,1,-10)	0.9982	0.9986	1.0000*	0.9999	0.9971	0.9999	1.0000*
Scale	SN(0,1.05,-10)	0.1678*	0.0379	0.1092	0.0833	0.1125	0.0983	0.1083
	SN(0,1.1,-10)	0.4808*	0.0782	0.3060	0.2299	0.3094	0.2694	0.3226
	SN(0,1.15,-10)	0.8092*	0.1475	0.5877	0.4770	0.5840	0.5409	0.6171
	SN(0,1.2,-10)	0.9621*	0.2450	0.8324	0.7446	0.8259	0.7959	0.8551
	SN(0,1.3,-10)	0.9998*	0.4020	0.9875	0.9780	0.9869	0.9839	0.9945
	SN(0,1.4,-10)	1.0000*	0.5154	0.9998	0.9995	0.9995	0.9995	1.0000*
	SN(0,1.5,-10)	1.0000	0.5705	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	St(0,1,-10,500)	0.0786*	0.0719	0.0458	0.0424	0.0416	0.0426	0.0521
	St(0,1,-10,100)	0.1039*	0.0549	0.0423	0.0395	0.0396	0.0392	0.0475
	St(0,1,-10,10)	0.6697*	0.0102	0.1393	0.0717	0.1016	0.0869	0.1159
	St(0,1,-10,7)	0.8914*	0.0097	0.2282	0.1139	0.1597	0.1351	0.2075
	St(0,1,-10,5)	0.9896*	0.0061	0.3702	0.2182	0.2822	0.2542	0.4328
	St(0,1,-10,4)	0.9991*	0.0058	0.5064	0.3759	0.4223	0.4091	0.6682
	St(0,1,-10,3)	1.0000	0.0023	0.7115	0.6794	0.6572	0.6736	0.9415

*最高檢定力、**異常值、***信心水準

此表為第四章第二節，在樣本數 500 之下模擬 10000 次的偏斜常態分配 KLIC 與 KS 統計量之型一誤差率及檢定力。大致具有一致性，除了 KLIC2 相反之外，其他六統計量之檢定力皆隨著分配差異越大而上升，同樣厚尾分配表現略遜於散佈與位置參數變化。

表 3 樣本數 500 時，Fréchet 分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0307	0.0309	0.0248	0.1589	0.1589	0.1579	0.0860
α^{**}	Fréchet(0,1,5)	0.0506	0.0476	0.0496	0.0490	0.0501	0.0485	0.0438
Location	Fréchet(0.02,1,5)	0.0748	0.1665	0.1542	0.1780	0.1732	0.1795	0.2054*
	Fréchet(0.03,1,5)	0.1564	0.3089	0.3263	0.3661	0.3344	0.3592	0.4169*
	Fréchet(0.05,1,5)	0.5589	0.7246	0.8183	0.8330	0.7635	0.8115	0.8774*
	Fréchet(0.08,1,5)	0.9836	0.9879	0.9993	0.9990	0.9957	0.9989	0.9997*
	Fréchet(0.1,1,5)	1.0000	0.9997	1.0000	1.0000	0.9999	1.0000	1.0000
	Fréchet(0.15,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0.2,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Scale	Fréchet(0,1.02,5)	0.0845	0.1136	0.1290	0.1406	0.1581	0.1530	0.1931*
	Fréchet(0,1.05,5)	0.5128	0.5514	0.7165	0.7340	0.7204	0.7396	0.8443*
	Fréchet(0,1.1,5)	0.9986	0.9915	0.9999	0.9999	0.9998	1.0000*	1.0000*
	Fréchet(0,1.15,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.2,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.25,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.3,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	Fréchet(0,1,4.8)	0.1193*	0.0336	0.0762	0.0576	0.0864	0.0712	0.0512
	Fréchet(0,1,4.5)	0.4416*	0.0569	0.2215	0.1882	0.2956	0.2475	0.1085
	Fréchet(0,1,4)	0.9723*	0.2364	0.8125	0.7740	0.8775	0.8444	0.4507
	Fréchet(0,1,3.5)	1.0000*	0.6422	0.9986	0.9979	0.9994	0.9993	0.9410
	Fréchet(0,1,3)	0.9999	0.9342	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1,2.5)	0.9998	0.9975	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1,2)	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

*最高檢定力、**信心水準

此表為第四章第二節，在樣本數 500 之下模擬 10000 次的右偏(Fréchet)分配 KLIC 與 KS 統計量之型一誤差率及檢定力。七統計量皆有顯著的檢定力，隨分配差異越大而上升。其 KLIC1 最下方三格所產生的異常判斷應為積分範圍選取所導致，其誤差小於千分位數，可忽略不計。

表 4 樣本數 200 時，常態分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0513	0.0513	0.0505	0.2382	0.2380	0.2362	0.1350
α^*	N(0,1)	0.0511	0.0502	0.0511	0.0518	0.0516	0.0507	0.0428
Location	N(0.1,1)	0.1053	0.1059	0.1093	0.1025	0.1044	0.1039	0.2441*
	N(0.2,1)	0.2991	0.3006	0.3093	0.2912	0.2939	0.2978	0.7480*
	N(0.3,1)	0.6212	0.6226	0.6380	0.5968	0.5971	0.6175	0.9804*
	N(0.4,1)	0.8906	0.8897	0.9000	0.8657	0.8663	0.8823	1.0000*
	N(0.5,1)	0.9837	0.9835	0.9852	0.9766	0.9756	0.9808	1.0000*
	N(0.7,1)	1.0000	0.9999	1.0000	0.9998	0.9997	0.9998	1.0000
	N(1,1)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Scale	N(0,1.1)	0.1397	0.1746	0.1644	0.0736	0.1926*	0.1332	0.0554
	N(0,1.2)	0.4807	0.5263	0.5233	0.2399	0.5589*	0.4249	0.1023
	N(0,1.3)	0.8395	0.8567	0.8643	0.5341	0.8770*	0.7782	0.2218
	N(0,1.4)	0.9736	0.9769	0.9791	0.8010	0.9804*	0.9513	0.3864
	N(0,1.5)	0.9980	0.9981	0.9982	0.9416	0.9985*	0.9957	0.6002
	N(0,2)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9977
	N(0,3)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	t(500,0)	0.0362	0.0327	0.0151	0.0301	0.0293	0.0311	0.0402*
	t(100,0)	0.0426*	0.0275	0.0168	0.0303	0.0308	0.0310	0.0408
	t(10,0)	0.2452*	0.0130	0.0296	0.0375	0.0470	0.0418	0.0408
	t(7,0)	0.4280*	0.0135	0.0391	0.0465	0.0569	0.0526	0.0464
	t(4,0)	0.8631*	0.0103	0.0788	0.0734	0.0971	0.0866	0.0565
	t(2,0)	0.9988*	0.0065	0.2351	0.3278	0.2467	0.2898	0.1571
	t(1,0)	0.8383	0.1859	0.8497	0.9647*	0.8213	0.9484	0.8517

*最高檢定力、**信心水準

此表為第四章第二節，在樣本數 200 之下模擬 10000 次的常態分配 KLIC 與 KS 統計量之型一誤差率及檢定力。大致具有一致性，隨著分配差異越大而上升，在厚尾分配表現上較散佈與位置參數變動來的不穩定。

表 5 樣本數 200 時，(左)偏斜常態分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0478	0.0494	0.0340	0.2053	0.2064	0.2049	0.1350
α^{**}	SN(0,1,-10)	0.0535	0.0504	0.0524	0.0545	0.0533	0.0549	0.0398
Location	SN(-0.01,1,-10)	0.0483	0.0573*	0.0521	0.0564	0.0552	0.0566	0.0433
	SN(-0.02,1,-10)	0.0479	0.0683*	0.0592	0.0599	0.0634	0.0618	0.0444
	SN(-0.03,1,-10)	0.0440	0.0747*	0.0676	0.0638	0.0724	0.0670	0.0560
	SN(-0.05,1,-10)	0.0462	0.1145*	0.0976	0.0914	0.0992	0.0961	0.0942
	SN(-0.08,1,-10)	0.0758	0.2106	0.1936	0.1682	0.1659	0.1699	0.1978*
	SN(-0.1,1,-10)	0.1128	0.2967	0.3021	0.2698	0.2447	0.2654	0.3094*
	SN(-0.2,1,-10)	0.6954	0.8072	0.9361*	0.8890	0.7654	0.8479	0.9231
Scale	SN(0,1.05,-10)	0.1259*	0.0295	0.0715	0.0654	0.0814	0.0754	0.0623
	SN(0,1.1,-10)	0.2782*	0.0305	0.1377	0.1164	0.1479	0.1337	0.1270
	SN(0,1.15,-10)	0.4874*	0.0442	0.2495	0.2236	0.2662	0.2511	0.2596
	SN(0,1.2,-10)	0.6923*	0.0662	0.3936	0.3533	0.3962	0.3833	0.4155
	SN(0,1.3,-10)	0.9360*	0.0993	0.6723	0.6618	0.6714	0.6764	0.7398
	SN(0,1.4,-10)	0.9940*	0.1285	0.8718	0.8793	0.8533	0.8770	0.9307
	SN(0,1.5,-10)	0.9999*	0.1427	0.9591	0.9699	0.9448	0.9629	0.9877
Shape	St(0,1,-10,500)	0.0800*	0.0693	0.0467	0.0459	0.0455	0.0445	0.0392
	St(0,1,-10,100)	0.0923*	0.0530	0.0462	0.0419	0.0429	0.0422	0.0384
	St(0,1,-10,10)	0.3761*	0.0145	0.0688	0.0536	0.0592	0.0567	0.0628
	St(0,1,-10,7)	0.5686*	0.0085	0.0871	0.0756	0.0724	0.0747	0.0836
	St(0,1,-10,5)	0.7845*	0.0045	0.1215	0.1052	0.0995	0.1024	0.1403
	St(0,1,-10,4)	0.9034*	0.0028	0.1575	0.1503	0.1267	0.1372	0.2140
	St(0,1,-10,3)	0.9808*	0.0021	0.2371	0.2715	0.1898	0.2294	0.4057

*最高檢定力、**信心水準

此表為第四章第二節，在樣本數 200 之下模擬 10000 次的偏斜常態分配 KLIC 與 KS 統計量之型一誤差率及檢定力。大致具有一致性，除 KLIC2 之外，其他六統計量之檢定力皆隨分配差異越大而上升，整體而言在厚尾分配表現上除了 KLIC1，其他皆略遜於散佈與位置參數變動。

表 6 樣本數 200 時，Fréchet 分配情況之檢定力

Test Statistics		KLIC1	KLIC2	KLIC3	KLIC4	KLIC5	KLIC6	KS
95% Critical Value		0.0596	0.0600	0.0437	0.2227	0.2239	0.2218	0.1350
α^{**}	Fréchet(0,1,5)	0.0465	0.0530	0.0556	0.0555	0.0517	0.0532	0.0396
Location	Fréchet(0.02,1,5)	0.0515	0.0963	0.0918	0.0886	0.0969*	0.0941	0.0906
	Fréchet(0.03,1,5)	0.0688	0.1453	0.1424	0.1518	0.1475	0.1553	0.1733*
	Fréchet(0.05,1,5)	0.1784	0.3276	0.3946	0.4054	0.3620	0.3944	0.4524*
	Fréchet(0.08,1,5)	0.5763	0.7010	0.8477	0.8479	0.7476	0.8190	0.8819*
	Fréchet(0.1,1,5)	0.8555	0.8753	0.9766	0.9745	0.9238	0.9622	0.9812*
	Fréchet(0.15,1,5)	0.9994	0.9966	1.0000	1.0000	0.9993	1.0000	1.0000
	Fréchet(0.2,1,5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Scale	Fréchet(0,1.02,5)	0.0643	0.0788	0.0916	0.0880	0.1012	0.0940	0.1018*
	Fréchet(0,1.05,5)	0.2163	0.2285	0.3381	0.3489	0.3514	0.3571	0.4489*
	Fréchet(0,1.1,5)	0.7951	0.6931	0.9197	0.9227	0.8780	0.9116	0.9623*
	Fréchet(0,1.15,5)	0.9952	0.9420	0.9990	0.9990	0.9953	0.9981	1.0000*
	Fréchet(0,1.2,5)	1.0000	0.9920	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.25,5)	1.0000	0.9989	1.0000	1.0000	1.0000	1.0000	1.0000
	Fréchet(0,1.3,5)	1.0000	0.9995	1.0000	1.0000	1.0000	1.0000	1.0000
Shape	Fréchet(0,1,4.8)	0.0931*	0.0306	0.0655	0.0559	0.0712	0.0629	0.0444
	Fréchet(0,1,4.5)	0.2461*	0.0246	0.1076	0.0953	0.1316	0.1166	0.0580
	Fréchet(0,1,4)	0.7142*	0.0470	0.3541	0.3511	0.4173	0.3947	0.1594
	Fréchet(0,1,3.5)	0.9760*	0.1231	0.7883	0.7950	0.8222	0.8221	0.4430
	Fréchet(0,1,3)	0.9994*	0.3111	0.9893	0.9885	0.9900	0.9920	0.8416
	Fréchet(0,1,2.5)	1.0000*	0.6225	0.9999	0.9999	0.9999	0.9999	0.9921
	Fréchet(0,1,2)	1.0000	0.8990	1.0000	1.0000	1.0000	1.0000	0.9999

*最高檢定力、**信心水準

此表為第四章第二節，在樣本數 200 之下模擬 10000 次的 Fréchet 分配 KLIC 與 KS 統計量之型一誤差率及檢定力。大致具有一致性，雖其檢定力不比樣本數 500、1000 來的高，但相對其他兩分配來的規律許多。

表 7 樣本數 2000 之配對前後各統計量比較

Test Statistics		KLIC1		KLIC2		KLIC3		KLIC4		KLIC5		KLIC6		KS		Proportion of matched No.	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
ln 2	Before	0.2815	0.0009	0.2480	0.0007	0.2648	0.0008	0.6561	0.0013	0.5969	0.0011	0.6265	0.0011	0.3074	0.0005	0.9831	0.0002
	After	0.0005	0.0000	0.0006	0.0000	0.0005	0.0000	0.0082	0.0001	0.0079	0.0001	0.0081	0.0001	0.0142	0.0001		
ln 3	Before	0.5922	0.0012	0.4893	0.0009	0.5408	0.0010	1.0004	0.0014	0.8891	0.0011	0.9447	0.0011	0.4309	0.0004	0.8720	0.0005
	After	0.0006	0.0000	0.0010	0.0000	0.0008	0.0000	0.0202	0.0001	0.0195	0.0001	0.0199	0.0001	0.0258	0.0001		
ln 4	Before	0.8021	0.0013	0.6483	0.0009	0.7252	0.0010	1.1994	0.0013	1.0541	0.0011	1.1268	0.0010	0.4989	0.0004	0.7531	0.0005
	After	0.0009	0.0000	0.0016	0.0000	0.0012	0.0000	0.0311	0.0002	0.0303	0.0001	0.0307	0.0001	0.0344	0.0001		
ln 5	Before	0.9365	0.0012	0.7616	0.0010	0.8490	0.0010	1.3210	0.0013	1.1630	0.0011	1.2420	0.0010	0.5429	0.0004	0.6659	0.0005
	After	0.0012	0.0000	0.0023	0.0000	0.0018	0.0000	0.0402	0.0002	0.0395	0.0002	0.0398	0.0002	0.0409	0.0002		
ln 6	Before	1.0271	0.0012	0.8486	0.0010	0.9379	0.0010	1.3989	0.0012	1.2418	0.0012	1.3203	0.0009	0.5740	0.0004	0.6029	0.0005
	After	0.0015	0.0000	0.0031	0.0000	0.0023	0.0000	0.0482	0.0002	0.0474	0.0002	0.0478	0.0002	0.0462	0.0002		
ln 7	Before	1.0905	0.0011	0.9190	0.0011	1.0048	0.0010	1.4519	0.0012	1.3034	0.0012	1.3777	0.0009	0.5972	0.0004	0.5560	0.0005
	After	0.0019	0.0000	0.0040	0.0000	0.0029	0.0000	0.0553	0.0002	0.0547	0.0002	0.0550	0.0002	0.0509	0.0002		
ln 8	Before	1.1378	0.0011	0.9775	0.0011	1.0576	0.0010	1.4901	0.0012	1.3536	0.0012	1.4219	0.0009	0.6156	0.0003	0.5194	0.0004
	After	0.0021	0.0000	0.0048	0.0000	0.0035	0.0000	0.0613	0.0002	0.0607	0.0002	0.0610	0.0002	0.0548	0.0002		
ln 9	Before	1.1750	0.0011	1.0283	0.0012	1.1016	0.0010	1.5195	0.0012	1.3965	0.0013	1.4580	0.0009	0.6308	0.0003	0.4900	0.0004
	After	0.0025	0.0000	0.0056	0.0000	0.0040	0.0000	0.0671	0.0003	0.0667	0.0003	0.0669	0.0003	0.0586	0.0002		
ln 10	Before	1.2055	0.0011	1.0722	0.0012	1.1389	0.0010	1.5437	0.0012	1.4334	0.0013	1.4885	0.0009	0.6436	0.0003	0.4663	0.0004
	After	0.0028	0.0000	0.0064	0.0000	0.0046	0.0000	0.0722	0.0003	0.0719	0.0003	0.0720	0.0003	0.0617	0.0002		

此表為第四章第三節總樣本數 2000 的配對前後七統計量與樣本比例資料。最左排為變量之係數，由上而下表示兩組重疊比例由高到低，網底為配對前，最右排為配對後實驗組個數除以配對前實驗組個數，其中 μ 為平均數、 σ 為標準差。整體而言統計量與配對比例呈現負相關。

表 8 樣本數 1000 之配對前後各統計量比較

Test Statistics		KLIC1		KLIC2		KLIC3		KLIC4		KLIC5		KLIC6		KS		Proportion of matched No.	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
ln 2	Before	0.2903	0.0013	0.2486	0.0010	0.2695	0.0011	0.6659	0.0018	0.5932	0.0016	0.6295	0.0016	0.3178	0.0007	0.9773	0.0003
	After	0.0005	0.0000	0.0008	0.0000	0.0007	0.0000	0.0107	0.0001	0.0103	0.0001	0.0105	0.0001	0.0221	0.0001		
ln 3	Before	0.5951	0.0018	0.4744	0.0012	0.5347	0.0014	1.0027	0.0019	0.8697	0.0015	0.9362	0.0015	0.4386	0.0006	0.8680	0.0007
	After	0.0007	0.0000	0.0012	0.0000	0.0010	0.0000	0.0223	0.0002	0.0215	0.0002	0.0219	0.0002	0.0339	0.0002		
ln 4	Before	0.7935	0.0018	0.6232	0.0012	0.7083	0.0014	1.1908	0.0018	1.0258	0.0015	1.1083	0.0014	0.5056	0.0006	0.7503	0.0008
	After	0.0009	0.0000	0.0019	0.0000	0.0014	0.0000	0.0327	0.0002	0.0317	0.0002	0.0322	0.0002	0.0431	0.0002		
ln 5	Before	0.9173	0.0017	0.7308	0.0013	0.8240	0.0014	1.3012	0.0017	1.1294	0.0016	1.2153	0.0013	0.5485	0.0005	0.6641	0.0007
	After	0.0011	0.0000	0.0027	0.0000	0.0019	0.0000	0.0415	0.0002	0.0404	0.0002	0.0409	0.0002	0.0501	0.0002		
ln 6	Before	1.0011	0.0016	0.8150	0.0014	0.9080	0.0014	1.3726	0.0017	1.2061	0.0016	1.2894	0.0013	0.5794	0.0005	0.6013	0.0007
	After	0.0014	0.0000	0.0035	0.0000	0.0025	0.0000	0.0492	0.0003	0.0481	0.0003	0.0486	0.0003	0.0559	0.0002		
ln 7	Before	1.0595	0.0016	0.8825	0.0015	0.9710	0.0014	1.4207	0.0016	1.2651	0.0017	1.3429	0.0012	0.6022	0.0005	0.5542	0.0007
	After	0.0017	0.0000	0.0044	0.0000	0.0030	0.0000	0.0561	0.0003	0.0550	0.0003	0.0555	0.0003	0.0609	0.0003		
ln 8	Before	1.1019	0.0015	0.9402	0.0015	1.0211	0.0014	1.4545	0.0016	1.3157	0.0017	1.3851	0.0012	0.6208	0.0005	0.5180	0.0006
	After	0.0019	0.0000	0.0052	0.0000	0.0036	0.0000	0.0617	0.0003	0.0607	0.0003	0.0612	0.0003	0.0648	0.0003		
ln 9	Before	1.1385	0.0015	0.9891	0.0016	1.0638	0.0014	1.4830	0.0016	1.3559	0.0017	1.4194	0.0012	0.6359	0.0005	0.4884	0.0006
	After	0.0022	0.0000	0.0061	0.0000	0.0041	0.0000	0.0675	0.0003	0.0665	0.0003	0.0670	0.0003	0.0689	0.0003		
ln 10	Before	1.1658	0.0015	1.0329	0.0017	1.0993	0.0014	1.5037	0.0015	1.3931	0.0018	1.4484	0.0012	0.6483	0.0005	0.4651	0.0006
	After	0.0024	0.0000	0.0069	0.0001	0.0047	0.0000	0.0724	0.0004	0.0714	0.0004	0.0719	0.0004	0.0725	0.0003		

此表為第四章第三節總樣本數 1000 的配對前後七統計量與樣本比例資料。最左排為變量之係數，由上而下表示兩組重疊比例由高到低，網底為配對前，最右排為配對後實驗組個數除以配對前實驗組個數，其中 μ 為平均數、 σ 為標準差。整體而言統計量與配對比例呈現負相關。

表 9 樣本數 500 之配對前後各統計量比較

Test Statistics		KLIC1		KLIC2		KLIC3		KLIC4		KLIC5		KLIC6		KS		Proportion of matched No.	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
ln 2	Before	0.3040	0.0019	0.2524	0.0014	0.2782	0.0016	0.6790	0.0026	0.5908	0.0021	0.6349	0.0022	0.3328	0.0010	0.9678	0.0005
	After	0.0007	0.0000	0.0011	0.0000	0.0009	0.0000	0.0148	0.0001	0.0143	0.0001	0.0146	0.0001	0.0358	0.0002		
ln 3	Before	0.6006	0.0025	0.4615	0.0016	0.5310	0.0019	1.0070	0.0027	0.8493	0.0020	0.9282	0.0021	0.4500	0.0008	0.8594	0.0010
	After	0.0008	0.0000	0.0017	0.0000	0.0012	0.0000	0.0258	0.0002	0.0249	0.0002	0.0254	0.0002	0.0465	0.0003		
ln 4	Before	0.7853	0.0026	0.6006	0.0017	0.6930	0.0020	1.1815	0.0026	0.9974	0.0021	1.0894	0.0020	0.5151	0.0008	0.7454	0.0011
	After	0.0010	0.0000	0.0024	0.0000	0.0017	0.0000	0.0359	0.0003	0.0346	0.0003	0.0353	0.0003	0.0560	0.0003		
ln 5	Before	0.9009	0.0024	0.7036	0.0019	0.8023	0.0019	1.2837	0.0024	1.0972	0.0021	1.1904	0.0019	0.5577	0.0007	0.6593	0.0010
	After	0.0012	0.0000	0.0033	0.0000	0.0022	0.0000	0.0443	0.0003	0.0429	0.0003	0.0436	0.0003	0.0636	0.0003		
ln 6	Before	0.9746	0.0023	0.7829	0.0019	0.8787	0.0019	1.3458	0.0023	1.1703	0.0022	1.2581	0.0018	0.5875	0.0007	0.5983	0.0010
	After	0.0013	0.0000	0.0042	0.0000	0.0028	0.0000	0.0513	0.0003	0.0498	0.0003	0.0506	0.0003	0.0693	0.0003		
ln 7	Before	1.0307	0.0022	0.8515	0.0021	0.9411	0.0019	1.3913	0.0022	1.2310	0.0023	1.3112	0.0018	0.6114	0.0007	0.5501	0.0009
	After	0.0015	0.0000	0.0052	0.0001	0.0034	0.0000	0.0582	0.0004	0.0567	0.0004	0.0575	0.0004	0.0752	0.0004		
ln 8	Before	1.0696	0.0022	0.9045	0.0021	0.9870	0.0019	1.4221	0.0022	1.2755	0.0023	1.3488	0.0018	0.6282	0.0007	0.5144	0.0009
	After	0.0018	0.0001	0.0061	0.0001	0.0039	0.0001	0.0642	0.0004	0.0626	0.0004	0.0634	0.0004	0.0799	0.0004		
ln 9	Before	1.1027	0.0021	0.9552	0.0023	1.0289	0.0020	1.4465	0.0022	1.3200	0.0025	1.3832	0.0018	0.6434	0.0007	0.4857	0.0009
	After	0.0020	0.0001	0.0070	0.0001	0.0045	0.0001	0.0693	0.0004	0.0677	0.0004	0.0685	0.0004	0.0836	0.0004		
ln 10	Before	1.1298	0.0021	0.9977	0.0023	1.0638	0.0020	1.4668	0.0022	1.3557	0.0025	1.4112	0.0017	0.6560	0.0006	0.4618	0.0008
	After	0.0021	0.0001	0.0079	0.0001	0.0050	0.0001	0.0739	0.0005	0.0724	0.0005	0.0732	0.0005	0.0878	0.0004		

此表為第四章第三節總樣本數 500 的配對前後七統計量與樣本比例資料。最左排為變量之係數，由上而下表示兩組重疊比例由高到低，網底為配對前，最右排為配對後實驗組個數除以配對前實驗組個數，其中 μ 為平均數、 σ 為標準差。整體而言統計量與配對比例呈現負相關。