

國立交通大學
運輸科技與管理學系碩士班

碩士論文



旅次分配推估方法之研究

研究生：李亦晴

指導教授：卓訓榮 教授

中華民國九十八年七月

旅次分配推估方法之研究
Estimation Of Trip Distribution

研究生:李亦晴

Student:Yi-cing Lee

指導教授:卓訓榮

Advisor:Hsun-Jung Cho

國立交通大學

運輸科技與管理研究所



Submitted to Institute of Transportation Technology and Management

College of Management

National Chiao Tung University

In

July 2009

中華民國九十八年七月

旅次分佈推估方法之研究

研究生:李亦晴

指導教授:卓訓榮

國立交通大學

運輸科技與管理研究所



運輸規劃中，交通起迄流量是基本且重要的資料，傳統的調查方法大多費時耗力且成本高，然而在交通流量大時，進行母體調查並不容易；本研究目的在利用可獲得資料推估交通旅次起迄資料，交通量的產生源自於運輸需求，然而，運輸需求是一種衍生需求，意即運輸並非旅行者真正目的，旅行者是為達成某種社會經濟活動而產生對運輸的需求；在古典引力模式中，僅探討可量測的變項，如：人口分佈、相異兩質點的距離，但是，交通旅次產生的潛在需求是不易被量測的，更重要的是，這種不易被量測的潛在需求正是交通量產生之泉源。本研究嘗試利用數學式來描述這些抽象的變項，建構一個可推估旅次資料的引力模式。

本研究亦對以路段資訊反推高速公路起迄表之可行性進行研究探討，構建以路段資訊推估高速公路旅次起迄分佈矩陣之模式，並以國道一號沿線北自汐止、南至岡山的電子收站資料作模式驗證，以瞭解模式之可行性，希望藉此避免傳統調查方法之困難又可獲得可靠的高速公路車輛起迄表。

Estimation Of Trip Distribution

Student: Yi-cing Lee

Advisor: Hsun-Jung Cho

National Chiao Tung University

Institute of Transportation Technology and Management

Abstract

In transportation planning, traffic flow of origin-destination is an vital and basic data. Traditional survey method expend a large number of labor power and material resources, the more important thing is that it's not easy to survey population while the traffic flow is large. The aim of this study is to modify the classical gravity model, and to use available data estimating origin-destination trip matrices. Traffic is based on transportation demand, however, transportation demand is a derived demand. It means that transportation is not the main purpose of traveler, the traveler requires transportation because of they want to achieve some economics activity. In classical gravity model ,it merely discuss the measurable variable, it doesn't care about the unmeasurable variable, like population , distance between origin and destination. But, the potential demand producted by transportation is uneasy to be measure, the more critical is that the potential demand it is a source of traffic. This study use mathematical-type to describe the abstract variables, and provide a model for estimating highway trip matrices. Another, this study develop a model for estimating based on link information as well. Finally, using available data to realize the feasibility of the model in this study.

誌謝

感謝眾人的支持與協助，得以完成本文。首先，感謝恩師 卓訓榮教授，在這兩年來亦步亦趨的指導與訓練，使得學生在研究上能夠有小小的成果；並且在邏輯思考層面上，獲得此生受用不盡的概念。論文口試之際，感謝韓復華教授與周幼珍教授，在百忙之中撥冗細審，並提供諸多建議與指教，使本論文漸臻完備，在此表謝忱。

感謝實驗室的所有成員，謝謝昱光學長、健綸學長在研究以及人生觀上面的指導與分享；謝謝黃恆學長、日錦學長、嘉駿學長在執行雷達計畫案的協助與共患難；謝謝如君陪伴我研究所這兩年，以及一同前往北京當交換學生互相扶持的日子，謝謝同學老總、猴給、建嘉在學業上給予我的支援，以及生活上的協助，你們都是我研究生活上不可或缺的同伴；感謝學弟志霖、學妹怡婷，因為有你們的加入，讓實驗室更顯活耀；感謝已畢業的學長姐：宜珊、之音以及小捲，雖然只共處了一年，但因為你們的關心與熱心，讓碩一甫進實驗室的我備感溫馨。

感謝在新竹的朋友，永泰、阿邦、冬瓜、阿振、筱雯，因為有你們，讓我在新竹的日子能夠精采有趣，特別是永泰在研究程式上不厭其煩的講解，因為你的協助與支持，使得我得以順利畢業。另外，還有我的高中同學們，對初到新竹的我照顧有加，以及大學同學們，尤其是球，在閒暇之餘互相關心、互吐苦水。

最後，感謝父母二十幾年來一路上無私的付出，因為你們的支持，不論我何時能畢業，只為完成我去北京交換學生想法，得以無憂的學習與拓展視野；因為你們的支持，而得以在無慮的情況下完成學業；因為你們的支持，得以一路上選擇我自己想走的路，即便有時多轉了幾個彎兒，你們總是說「沒關係！」，簡短的一句話，囊括無限的包容，這對我而言，絕對是莫大的支持與動力，使我能繼續向前；還有哥哥的不具體卻是心中很踏實的支持，使我得以完成學位。此時此刻的我，是滿足而開心的，謹以此文獻給我的家人，因為你們對我的全力支持與付出而完成此文。衷心感謝！

李亦晴 謹誌

中華民國九十八年七月 于新竹

目錄

中文摘要.....	ii
英文摘要.....	iii
誌謝.....	iv
表目錄.....	viii
圖目錄.....	ix
第一章 緒論.....	1
1.1 研究動機	1
1.2 研究目的	2
1.3 研究方法及流程	3
第二章 文獻回顧.....	4
2.1 傳統車輛起迄調查方法之回顧	4
2.2 車輛起迄推估模式	6
2.2.1 極大熵模式	6
2.2.2 引力模式	7
2.2.3 干預機會模式	10
2.2.4 引力-機會模式	10
2.2.5 最大概似法	11
2.2.6 最小平方法	11
2.3 不完整資料之種類與處理	13
2.4 統計上處理不完整資料的方法	14
第三章 模式的建構.....	18
3.1 古典引力模式	18
3.2 廣義引力模式	20
3.2.1 空間分隔的量測	20
3.2.2 阻抗函數	22
3.2.3 定義引力模式	22
3.2.4 模式的原則	23

3.3 本研究模式	27
3.3.1 模式建構的原則	27
3.3.2 模式建構的基本假設	28
3.3.3 模式的特性	29
3.3.4 模式推導	32
3.3.4.1 模式(一)推導	33
3.3.4.2 模式(二)推導	34
第四章 抽樣分析與資料整理.....	38
4.1 抽樣方法	38
4.1.1 系統抽樣	38
4.2 檢定方法	40
4.3 資料蒐集	41
4.2.1 資料整理	42
4.2.2 抽樣比例與母體型態分析	44
第五章 模式推估與驗證.....	47
5.1 樣本與母體型態分析	47
5.2 起迄資料推估流程	48
5.3 模式優劣程度指標	49
5.4 模式推估與優劣程度測試	50
5.4.1 古典引力模式推估	50
5.4.2 模式(一).....	51
5.4.3 模式(二).....	52
5.4.4 小抽樣比例樣本推估	54
5.4.5 小結	56
第六章 不完整資料插補.....	58
6.1 引力模式插補不完整資料	58
6.2 不完整資料插補流程	59
6.3 不完整資料處理	60

第七章 結論與建議.....	61
7.1 結論.....	61
7.2 建議.....	62
參考文獻.....	63
附錄、一.....	65



表目錄

表 3-1	旅次起迄分佈矩陣表.....	24
表 4-2	車種代碼.....	42
表 4-4	方向.....	42
表 4-5	國道一號 交通車輛起迄表	44
表 4-6	每 2 輛車抽取第一輛車的起迄表	45
表 4-7	每 10 輛車抽取第一輛車的起迄表	45
表 4-8	每 20 輛車抽取第一輛車的起迄表	46
表 4-9	每 30 輛車抽取第一輛車的起迄表	46
表 5-1	模式間推估程度比較.....	56
表 5-2	模式間小抽樣比例樣本推估程度比較	57



圖目錄

圖 1-1 研究方法及流程.....	3
圖 2.1 都市區域圖.....	14
圖 3-1 子群體示意圖.....	27
圖 4-1 原始資料型態.....	41
圖 4-2 時間遞移抽樣 28 小時	43
圖 5-1 起迄資料推估流程圖	48
圖 5-3 模式(一)母體與推估起迄資料比較圖	51
圖 5-4 模式(二)母體與推估起迄資料比較圖	53
圖 5-5 古典引力:以抽樣比例小樣本推估資料與母體資料比較圖 ...	54
圖 5-6 模式(一):以抽樣比例小樣本推估資料與母體資料比較圖	55
圖 6-1 不完整資料插補流程圖	59
圖附-1 站 1 至站 2 間的路段流量.....	65
圖附-2 站 2 至站 3 間的路段流量	65
圖附-3 站 3 至站 4 間的路段流量	66

第一章 緒論

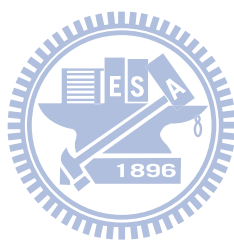
1.1 研究動機

高速公路車輛起迄表的資訊，傳統上有：路邊訪問、車輛牌照登錄、郵卡問卷、電話訪問、錄影或照相偵測、車輛標記、車輛追蹤、匝道收費、人口普查等等方法取得。但這些方法之資料收集過程各有其困難，資料的精確度也不同，更重要者是這些調查方法之功能亦不相同，為了達到某些特定目的，使得很多調查方法都有其侷限無法達成該目的。除了傳統調查方法之外，在某些較易取得的資訊下，是否有新的方法可以避免傳統調查之困難又可求得高速公路車輛起迄表？而這些方法之可行性且其限制又如何呢？期望構建一可推估高速公路旅次起迄分佈矩陣之模式，並以本研究調查之資料進行測試，以瞭解模式之可行性，希望藉此避免傳統調查方法之困難又可求得高速公路車輛起迄表。

。本研究首先針對傳統各起迄調查方法進行理論分析，亦即就高速公路之起迄特性，探討應該採用何種調查較佳及面臨的問題為何？其次就較適合高速公路起迄調查之車輛牌照登錄法中的攝影、錄音及人工抄錄三種方式進行實作分析，透過實際調查一方面將三種調查方式（攝影、錄音、人工抄錄）所遭遇到的困難及問題列出，另一方面也就高速公路匝道形式及匝道交通量等不同狀況，應該採用哪一種調查方式較佳作成建議，以供未來應用之參考。另外，本研究亦對以路段資訊反推高速公路起迄表之可行性進行研究探討，構建以路段資訊推估高速公路旅次起迄分佈矩陣之模式，並以本研究調查之資料進行測試，以瞭解模式之可行性，希望藉此避免傳統調查方法之困難又可求得高速公路車輛起迄表。（民 84,卓訓榮等）

1.2 研究目的

高速公路車輛起迄表是高速公路規劃、設計的主要參考依據，分別探討高速公路之各項資訊所推估起迄表模式，並以國道一號沿線北自汐止、南至岡山的電子收站資料作模式驗證，以探討其可行性；期望在資源有限的情況下，提出經濟且可靠的模式供高速公路車輛起迄推估，以及，以引力模式進行資料插補處理。



1.3 研究方法及流程

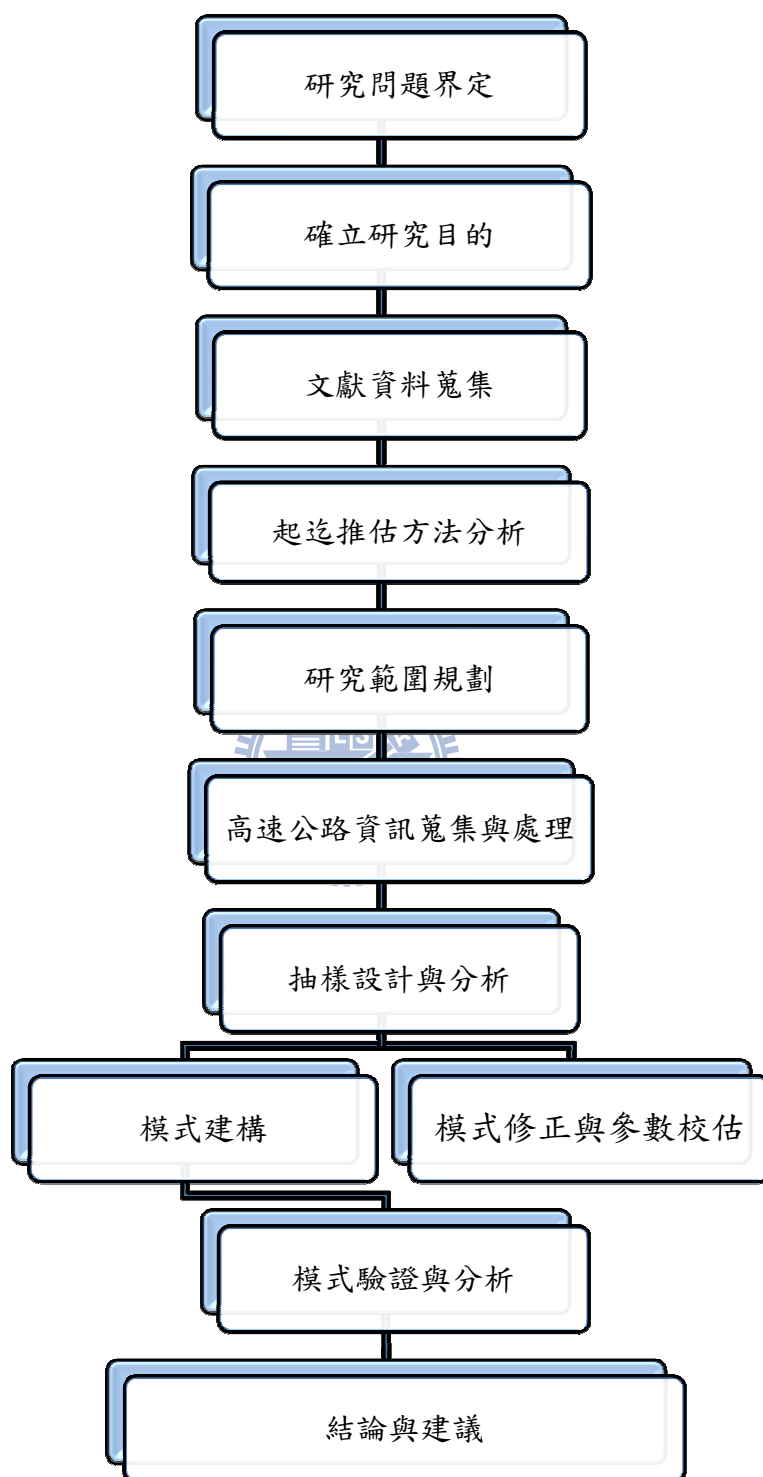


圖 1-1 研究方法及流程

[資料來源:本研究整理]

第二章 文獻回顧

2.1 傳統車輛起迄調查方法之回顧

1. 路邊訪問調查法

此法通常於規劃研究區之週界進行，在道路與週界相交之處設置調查站，由交通警察協助攔截通過之車輛，對其起迄點、目的等特性進行調查。將調查結果予以分析處理，資料分析時再以放大系數進行調整，即可獲得旅次分佈矩陣，但此法不適用於公速公路。

2. 車輛牌照登錄法

此法於車輛可能通過之地點設置調查站，將通過各調查站之車輛牌照號碼以人工、錄音或攝影方式登錄後，再將各站資料利用電腦進行比對，以得到旅次分佈矩陣。

該方法進行時不會干擾交通情況是其優點，但仍有一些缺點存在，如：

- (1) 車速高與車流量大時，錄音與人工抄錄調查員難以正確讀取及記錄牌照號碼，造成日後電腦作業上比對困難或產生錯誤。
- (2) 當一車輛之車號無法被尋獲配對時，將造成兩個以上之車號孤立或重覆車號，增加不準確率。
- (3) 當以攝影方式進行時則成本太高且拍攝亦受到天候影響而無法拍錄，且當攝影人員疏失時，將造成某些路口或匝道之車輛牌照全部無法辨識，則造成調查之全部資料無法運用。

3. 郵卡問卷法

此法即是於主要車輛經過地點，攔截車輛發給駕駛人已附回郵之調查卡，由駕駛人填妥，寄回研究單位，進行資料分析。該問卷不當場回收以免造成交通延誤，但也因此造成問卷回收率不高，無法反映實際情形。

4. 電話訪問法

以研究區內之住戶為對象，透過電話號碼抽樣進行，以電話訪問的方式調查其前一日之旅次資料。此法節省人力是其優點，其缺點為：

- (1) 僅包含家旅次，不包括非家旅次。
- (2) 調查之資料並無區外至區外通過性旅次。

5.車輛標記法

於調查站設立標誌指示駕駛人打開車燈，再於車流可能轉向之地點記錄打開車燈之車輛數以得到其轉至特定方向之車流。缺點在於駕駛人之配合度影響整個調查結果。

6.車輛追蹤法

以雇用車輛追蹤紀錄之方式進行研究區內的起迄點調查，但此方法實施困難且成本亦高。

7.匝道收費站法

收費公路之收費方式若為入匝道取票，出匝道付費者可直接統計其回收收票票證，以得車輛起迄點資料，此法在台灣之高速公路並不適用。

8.人口普查法

利用人口普查之時機，同時附帶調查交通特性，以獲得旅次起迄點之資料，此法若問卷詳細將可得到最精確的靜態資料作為基年資料之用。缺點是：

(1) 調查成本太高。

(2) 調查時間間隔太長，在快速成長地區較不適用。



2.2 車輛起迄推估模式

由於傳統起迄資料蒐集耗費能力物力及時間，因此，近年來，Ortuzar 和 Willumsen(2001)、Roy 和 Thill(2004)以及 Cascetta 等人(2007)，致力於利用簡易獲得的資訊、低成本、低人力的方式來推估旅次起迄方法。Wilson(1967)、Jayet(1990)將傳統的絕對空間距離概念轉換成人口遷移阻力，距離可作為具體測量中介機會。

常見的推估模型如:極大熵法(Entropy maximization)、引力模式(Gravity model)、干預機會模式(Intervening opportunities)、引力-機會模式(Gravity-opportunity model)以及最大概似法(Maximum likelihood)。以下茲作介紹:

2.2.1 極大熵模式

Clausius 的經驗準則-熱力學第二定律，極大熵的觀念是引用熱力學的觀念，將氣體分子結構中的分佈及穩定狀態，比擬為人在社會中的行為狀態，在各種限制件下找出最可能的系統狀態。該定律說明當一個熱力學系統達到最後熱平衡狀態時，該系統的熵會達到最大值；進一步的研究指出當系統的熵最大時，其自由能將成為最小。極大熵法是由 Wilson(1970)提出，主要改進了引力模式後得出；引力模式廣為運輸規劃者所使用，極大熵法亦受學者重視，兩者都需滿足旅次總產生數。在此一特性的影響下人們慣性的傾向於將熵視為類似能量的巨觀物理量；引力模式和極大熵法主要有別於，前者提出主要是物理觀念，後者是基於統計分析觀點，因此後者較具有推理之基礎。

將極大熵原理用來確定交通量分佈的基本想法，是把交通網路看作一個基本對稱系統，利用交通生成量及其分佈，定義系統狀態的概率分佈及對應於分佈的熵，然後在關於該系統狀態的約束條件下，求使熵最大的那個分佈，即所推估的交通量分佈。其模型如下:

概率分佈 T_{ij} 定義的熵為

$$H(P_{ij}, i, j = 1, 2, \dots, s) = -\sum_i \sum_j P_{ij} \ln P_{ij} \quad (2-1)$$

必須滿足為

$$\sum_i T_{ij} = O_i \quad (2-2)$$

$$\sum_j T_{ij} = D_j \quad (2-3)$$

利用最大熵原理，令

$$a_i = \frac{O_i}{T}, \quad b_j = \frac{D_j}{T}, \quad P_{ij} = \frac{T_{ij}}{T}$$

可知

$$\sum_i a_i = \sum_j b_j = 1 \quad (2-4)$$

$$\sum_j P_{ij} = a_i, \quad \sum_i P_{ij} = b_j \quad (2-5)$$

$$\sum_i \sum_j C_{ij} T_{ij} = C \quad (2-6)$$

在(2-1)(2-2)(2-3)限制式下求(2-1)極大值得

$$T_{ij} = KO_i D_j f(C_{ij}) \quad (2-7)$$

2.2.2 引力模式

引力模式早期是類比牛頓的萬有引力定律(式 2.1.1)，

$$F_{ij} = \gamma \frac{m_i m_j}{d_{ij}^2} \quad (2-8)$$

表示兩質量 m_i 和 m_j 在距離 d_{ij} 時萬有引力為 F_{ij} ； γ 為常數

而導出引力模式，Reilly(1931)將此概念應用到都市系統分析，其基本公式如下：

$$T_{ij} = k \frac{O_i D_j}{C_{ij}^2} \quad (2-9)$$

其中

T_{ij} : 由 i 至 j 的交通起迄量

O_i : 表示 i 點產生的交通量

D_j : 表示 j 點吸引的交通量

C_{ij} : 旅行成本

k : 為常數

由於在交通流量起迄表中必須滿足以下限制式：

$$\sum_i T_{ij} = O_i$$

$$\sum_j T_{ij} = D_j$$

故引力模式修正為以下一般式

$$T_{ij} = A_i B_j O_i D_j f(C_{ij}) \quad (2-10)$$

其中

$$A_i = \left[\sum_j B_j D_j f(C_{ij}) \right]^{-1}$$

$$B_j = \left[\sum_i A_i O_i f(C_{ij}) \right]^{-1}$$

以上兩式能滿足(2-2)(2-3)，以遞迴方式求解。

後來威爾生(Wilson, 1967)發現，引力模型可用最大熵法(entropy maximizing)導出，使得此模型更具推理基礎，且改稱此種模型為空間交互作用模型(spatial interaction model)。目前在規劃上引用最廣泛之都市空間模型者為勞利模型(lowry model)。推導引力模式以最大熵法之前，先介紹總成本限制式

$$\sum_i \sum_j T_{ij} C_{ij} = C \quad (2-11)$$

推導過程如下:

目標為極大化起迄矩陣 $\{T_{ij}\}$ 的熵分配

$$W(\{T_{ij}\}) = \frac{T!}{\prod_{ij} T_{ij}!}$$

T 為總起迄量

故可能產生的情況是 w ，且其數量超過 T_{ij}

$$w = \sum W(\{T_{ij}\})$$

限制式為

$$\sum_i T_{ij} = O_i$$

$$\sum_j T_{ij} = D_j$$

$$\sum_i \sum_j T_{ij} C_{ij} = C$$

以 Lagrangian multiplier 求解:

$$\mathcal{L} = \ln W + \sum_i \lambda_i (O_i - \sum_j T_{ij}) + \sum_j \lambda_j (D_j - \sum_i T_{ij}) + \beta (C - \sum_i \sum_j T_{ij} C_{ij})$$

$\lambda_i, \lambda_j, \beta$ 為 Lagrangian multipliers

因求解 $\ln W$ 比 W 有效率，故利用 Stirling 近似

$$\ln N! \approx N \ln N - N$$

則

$$\frac{\partial \ln N!}{\partial N} = \ln N$$

故

$$\text{令 } \frac{\partial \mathcal{L}}{\partial T_{ij}} = -\ln T_{ij} - \lambda_i - \lambda_j - \beta C_{ij} = 0$$

$$\ln T_{ij} = -\lambda_i - \lambda_j - \beta C_{ij}$$

$$T_{ij} = e^{-\lambda_i - \lambda_j - \beta C_{ij}}$$

故

$$T_{ij} = A_i B_j O_i D_j f(C_{ij})$$

各符號如前文所示，其中 $f(C_{ij})$ 為阻抗函數。

將引力模式發揚光大者為勞利模式(Lowry model)，但最初的勞利模型是一種潛能模型，用來處理人口的分佈。在這個模式內，配置任何一區的人口數量，取決於該區內各項分區潛能之總和。然而，加林(Garin,1966)加以修改，對相互作用所使用的度量則以引力模式為基礎，發展至今；Casey(1955)、Huff(1964)等人將他運用在起迄資訊推估。

2.2.3 干預機會模式

干預機會模式(Intervening opportunities)是假設一個旅行者滿足下一個旅行的機會是受到前一次旅行的影響。Stouffer(1940、1960)原始的干預機會模式是藉由 Wilson(1967、1970)極大熵模式推演而來，並且假定旅次與空間累積機率為線性關係(Oppengeim, 1980)，數學模式如下：

$$T_{ij} = k_i O_i [e^{-LV_{j-1}} - e^{-LV_j}] \quad (2-12)$$

其中 k_i 為校正因子， V_{j-1} 是前一區到 $j-1$ 區的累積機率，(2-12)必須滿足限制式

$$\sum_j T_{ij} = O_i \quad (2-13)$$

$$k_i = [1 - e^{-LV_n}]^{-1} \quad n \text{ 為總區域數，}$$

以 V_j 代替一般的 D_j ，意即吸引到 j 區的交通量相當於遷移到 j 區的真實機會(Stouffer,1940、1960)。

2.2.4 引力-機會模式

引力-機會模式(Gravity-opportunity model)是以干預機會代替成本的引力模式型態，將成本改寫為(Wills,1986、Roy,1933、Almeida and Goncalves,2001)，其基本公式如下：

$$C_{ij} = V_{j(i)} \quad (2-14)$$

雙限制式的引力-機會模式為

$$T_{ij} = a_i b_j O_i D_j \exp(-\beta V_{j(i)}) \quad (2-15)$$

$$a_i = \left[\sum_j b_j D_j \exp(-\beta V_{j(i)}) \right]^{-1}$$

$$b_j = \left[\sum_i a_i O_i \exp(-\beta V_{j(i)}) \right]^{-1}$$

其中 a_i 和 b_j 利用最大概似法則校估而來，(Goncalves and Ulysea Neto, 1993)；可獲得其結果類比於引力模型，(Evans, 1971)， a_i 和 b_j 為校正因子以確保流量限制式守恆。

2.2.5 最大概似法

由英國統計學家費雪(R.A Fisher, 1912)提出，最大概似法假設旅次分佈為 Poisson 分配，且已知路網中的路段旅次交通量，並以概似值來從中判斷最佳的 OD 矩陣。其模式一般形式為：

$$\text{Max} \quad \prod_{ij} \frac{\exp(-\rho_{ij} T_{ij}) (-\rho_{ij} T_{ij})^{t_{ij}}}{t_{ij}!}$$

s.t.

判斷準則為概似值越大時，該推估 OD 矩陣也越佳。

2.2.6 最小平方法

此法係利用先期起迄分佈矩陣及路段交通量來推估起迄分佈矩陣，且假設推估值與調查值具有不偏性，模式能提供估計的信賴區間，以供對於估計準確性的一個參考，但推估結果的正確與否與先期起迄分佈矩陣相關甚大。其模式一般型式為：

$$\begin{aligned} \text{Min} \quad & (\bar{T} - T)^T Z^{-1} (\bar{T} - T) + (\bar{f} - f)^T W^{-1} (\bar{f} - f) \\ \text{s.t.} \quad & f = PT \end{aligned}$$

其中

\bar{T} : 先期的旅次分佈矩陣

T : 估計的旅次分佈矩陣

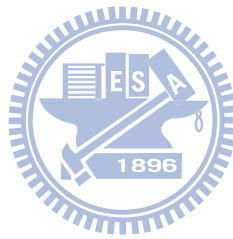
\bar{f} : 觀察的路段流量

f : 估計的路段流量

P : 起迄點使用特定路段的比例

Z : T 的協方差矩陣

W : f 的協方差矩陣



2.3 不完整資料之種類與處理

在一般調查中，資料不完整的種類有二：(1)某一抽樣單位(Sample unit)所有資料均未收集，此稱為 unit nonresponse；(2)某一種抽樣單位對某些問題不與回答而造成 item nonresponse。對於兩種不完整資料的處理通常是對於抽樣單位再度調查或以某種適合的方法代入缺失值(missing values)。缺失值可分為很多不同類型，不同型式的缺失值其處理方式不同，因此須先將旅次起迄分佈矩陣所產生的不完整資料予以分類，並根據分類原則定義本研究之問題範圍。本研究將旅次起迄分佈矩陣所產生不完整資料處理的情形分為以下四點：

- (1) 在每一高速公路匝道進出口採取車輛牌照登錄法，再進行車牌之比對，即可得到一組旅次起迄分佈矩陣。
- (2) 當我們針對都是區域進行家庭訪問等之抽樣調查，因為分區多，加上人力、財力不足，各區域選取的樣本很少，以致調查之結果可能沒有顯示出某些起迄分佈的旅次量，使得這些區域的旅次量因此被忽略，而產生不完整資料。
- (3) 調查之初，由於某一區域尚未開發而沒有旅次產生，經一段時間後，我們可以成長因素法則來求算未來地區旅次，但對於原本尚未開發區域則未加考慮，對於未來旅次量的估計會有偏差，便可能會有不完整資料產生。
- (4) 圖 2.1 表示將都市劃分為九個區域(1-9)，為節省調查費用，執行調查時僅對 1,2,4,5,7 的區域進行家庭訪問調查，及對河川(圖中以虛線表示)，實施周界調查。因此。我們之旅次分佈矩陣中將得不到 $3 \rightarrow 6, 8 \rightarrow 9, 9 \rightarrow 3, \dots$ ，此亦屬不完整資料。

1	2	3
4	5	6
7	8	9

圖 2.1 都市區域圖

(資料來源:Day & Hawkins,1979)

上述四種不完整資料其產生的原因皆異，後三種產生不完整資料的原因並非完全是因為實際旅次量少而造成；然而，第一種則純粹是因為實際的旅次量少，經抽樣使得微小的旅次量消失，或是，抽樣比例小造成不完整資料。本研究將針對第一種不完整資料型態的問題進行研究。

2.4 統計上處理不完整資料的方法

交通上處理不完整資料的方法多以交通上現有估計起迄分佈矩陣之模式為主，其缺點為它使用了目前尚未成熟的模式，或成熟的模式尚沒有適用於國內的參數值(指模式中所需的參數值)，因此本研究嘗試避開既有的模式，尋找統計上處理不完整模式的方法，分析各方法使用的限制與適用性。

進行隨機抽樣或系統抽樣而產生不完整資料時，在統計上的處理方式，大致包含下列三種(Aiffi and Elashodd,1966):

(1) 完整資料法(compelete-case analysis)

將所有有缺失值的資料剔除後，只分析完整的資料，其優點為簡單且能不經調整就適用於既有的統計軟體，缺點則為失去了部分資訊(loss of information)，使分析結果有偏差，尤其在有許多缺失值的情況。其相關研究包含 Keleinbaum,Margensten and Kupper(1981)及 Dixon(1983)等。

(2) 可用資料法(available-case analysis)

相對於只取完整資料法，可用資料法則只剔除缺失值分析所有可用資料，但也使得分析變得複雜。其相關研究包含 Dixon(1983)、Kim and Curry(1977)、Haitovsky(1968)、Azen and Van Guilder(1981)等。

(3) 插補法(fill-in method)

對每一缺失值代入適當的值，使現有資料成為完整資料，然後再用完整資料分析法進行分析。

上述三種方法中，前二種處理不完整資料的方式較為簡單，將不完整資料剔除即可，但旅次起迄分佈矩陣不似一般問卷調查，可將部份資料剔除再進行研究。它賴於起迄分佈表中每一起迄對之值皆存在，才得以成為一完整的旅次起迄分佈矩陣，進而再以此資料進行運輸規劃的研究。因此，我們必須採用插補法，將缺失值代入一適當的值，再用完整資料分析。下面僅就與本研究方法相關之插補法做較詳細的介紹。

在統計上，研究插補法的文獻很多，在其插補的分類方式亦隨分類精細度之不同，而有不同的分類方式，本研究擬以 Little and Rubin(1987)的分類原則簡介如下：

(1) 平均插補(Means imputation (filled in))

本方法是將觀察到的值與缺失值問題給予適當的權重，再將其值平均；亦或以尚存在完整值之平均值將缺失值補足。如 Buck's Method (Buck,1960)。

(2) 現期資料插補(Hot deck imputation)

首先觀察資料之分佈型態，對於其型態有所了解，再根據此分佈型態對缺失值部分給予一適當的經驗值。相關文獻包含 Ernsr(1980), Kalton and Kish(1981), Ford(1983), and David

et.al.(1986)

(3) 替代插補(Substitution imputation)

觀察變數間的相關性，對於高度相關的變數，其間之部分變數若存在著不完整資料，則可以其相關變數(完整資料)之值推導出缺失值的部分。

(4) 前期資料插補(Cold deck imputation)

本法純粹是以前期之資料作為根據，再以適當的插入。

(5) 迴歸插補(Regression imputation)

根據完整資料求取足以代表資料型態的迴歸線，再將缺失值之 x 值代入所得到之迴歸線，即可求出我們所須之 \hat{y} 。Buck's Method (Buck,1960)、Yate's Method(1933)、Bartlett(1937)皆是這方面的研究。

(6) 隨機迴歸插補(Stochastic regression imputation)

對於我們所得到之迴歸線 $y = a + bx + \varepsilon$ ，假設 ε 為 normal 或 binary，在本假設的基礎上建構處理不完整資料的模式。Herzog and Rubin (1983)研究過這方面的問題。

(7) 組合插補(Composite method)

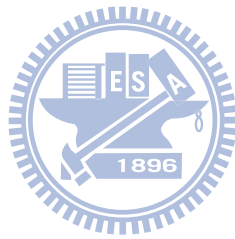
本方法是將前述之方法加以組合來處理不完整資料。如以 Hot deck+regression 這方面的問題於 Schieber(1978)中有詳細的討論。

(8) 多重插補(Multiple imputation method)

本方法是從已觀察得到的資料型態中，抽取多個樣本值，再將每一樣本值與其他觀察到得值做不同的組合，尋找一較適當的樣本值做為缺失值的插補。其缺點在於抽取的樣本值並不知為何種分配型態，也就是說，原始資料的型態與條件機率的型態(抽取之樣本值)可能並不相同；另外，此種方法須將各種樣本值與母體作不同的組合，一旦缺失值增加時，每一格

子點與其他格子點的組合更是繁多，使得資料處理的過程繁雜，是本方法的缺失。

統計上對於不完整資料的問題有多方面的討論，其方法中，大多需要資料的規律性，抑或需要較多額外的資料，當若干資料無法重新調查取得或沒有規律的存在時，很多方法將無法使用，因此本研究希望提出以較少的資料(即現有的資料)即可處理不完整資料的方法。對於 Little and Rubin(1987)之分類本研究再以其使用的工具予以分類，並說明應用於本研究之問題時可能遭遇的問題與限制所在。

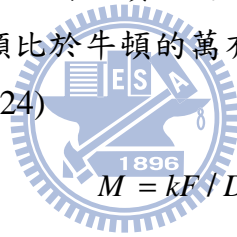


第三章 模式的建構

本研究所建構推估高速公路車輛起迄表的方法是以引力模式為基礎，並將此法加以改善，使其更適用於抽樣產生之車輛旅次起迄調查資料；並嘗試以引力模式作不完整資料插補。因此，以下將先詳盡的介紹引力模式，再根據不足之處予以改善，最後，在第五章以實際的例子來比較傳統引力模式與本研究所提出方法的優劣所在。

3.1 古典引力模式

引力模式早期是類比牛頓的萬有引力定律，該模式至今已有悠久的歷史，除了科學領域，目前已廣泛運用於各社會學領域。起初，觀察到人類的遷移行為可類比於牛頓的萬有引力定律，並將人類的行為以數學式表示: (Young,1924)



$$M = kF / D^2 \quad (3-1)$$

其中 k 為常數，M 是遷移的人口數，F 是區域吸引強度，D 是距離。

隨後，Reilly(1931)將引力模式應用在零售市場，描述消費者選擇店家的行為；Stewart(1941)提出人口統計引力模式(demographic gravitation)，假定人口中心 i 及 j 具有交互作用 T_{ij} ，定義為人口質量 P_i 和 P_j 的乘積，除以 i 和 j 距離平方，即

$$T_{ij} = GP_i P_j (d_{ij})^{-2} \quad G: \text{人口統計引力常數}, (3-2)$$

Stewart(1948、1950)、Dodd(1950)再予以權重(weights)關係，數學式如下:

$$T_{ij} = G(w_i P_i)(w_j P_j)(d_{ij})^{-2} \quad (3-3)$$

利用權重來表示人口質量的「異質性」；並以能量(demographic energy)來代替力量(demographic force)。

Huff(1963)在零售模型中，藉由消費者的集合探討選擇購物區位，假設 P_{ij} 為消費者自區位 i 至 j 的機率， t_{ij} 為 i 到 j 的旅行時間， $-\lambda$ 為引力指數， W_j 表示迄點 j 的相關引力，Huff 模型如下：

$$P_{ij} = \frac{W_j t_{ij}^{-\lambda}}{\sum_j W_j t_{ij}^{-\lambda}} \quad (3-4)$$

在 Huff 模型中，主要是移除了 Newtonian model 的嚴格限制：

- (1) 在兩個獨立個體中，需要處理的是旅行時間而非距離。
- (2) 允許引力參數(λ)從觀察值校估而來，而非直接指派為(-2)次方。

此外，關於原始的引力模式，未考慮到其他地區競爭的吸引，因此容易高估地區內之間的引力，修正後之引力模式可描述如下，在兩個或更多地區之間的相互作用量與分區大小(或吸引力)成正比，而就分區間之距離及參與競爭分區之間之相對吸引力成反比，並在模型中加入一校正因子以代表其他分區的競爭力，修正後的模型如下：

$$T_{ij} = P_i \frac{P_j d_{ij}^{-b}}{\sum_j P_j d_{ij}^{-b}} \quad (3-5)$$

其中 T_{ij} 表 i 區到 j 區的旅次，若

$$\text{令 } A_i = \left(\sum_j P_j d_{ij}^{-b} \right)^{-1},$$

上式將改寫為

$$T_{ij} = P_i A_i P_j d_{ij}^{-b},$$

意即 i 區到 j 區的或然率為 $A_i P_j d_{ij}^{-b}$ 。

3.2 廣義引力模式

本研究目的在利用可獲得資料推估交通旅次起迄資料，交通量的產生源自於運輸需求，然而，運輸需求是一種衍生需求，意即運輸並非旅行者真正目的，旅行者是為達成某種社會經濟活動而產生對運輸的需求；在古典引力模式中，僅探討可量測的變項，如：人口分佈、相異兩質點的距離，但是，交通旅次產生的潛在需求是不易被量測的，更重要的是，這種不易被量測的潛在需求正是交通量產生之泉源。以下，我們嘗試利用數學式來描述這些抽象的變項：

3.2.1 空間分隔的量測

Ashish Sen 和 Tony E. Smith(1995)假定：在相異兩質點當中，隱含數個空間分隔(spatial separation)，在行為者(actors)和機會(opportunity)中產生交互作用，這些交互作用受到個別行為者(如：通勤者、購物者)和機會(如：工作、商店各種經濟活動)的影響，在不同情況下，便產生相異的空間分隔。我們試圖了解行為者和機會所形成的空間分隔對交互作用的影響是吸引力或是排斥力，因此具體化上述變項；假定已知交互作用中，每個行為者集合 A ， $\alpha \in A$ ，機會集合 B ， $\beta \in B$ ，存在有限相關的分隔屬性(separation attributes)集合 $k \in K$ ，介於 A 和 B 之間，空間分隔可表示為有限集合 $\{c^k : k \in K\}$ ，存在分隔量(separation profile) $c(\alpha, \beta) = \{c^k(\alpha, \beta) : k \in K\}$ ，以 $c(\alpha, \beta)$ 來具體化表示對交互作用的影響所需的成本；進而以空間總合假設表達如下：

假設一組行為者形式 $\{\alpha_i : i \in I\}$ 和一組機會形式 $\{\beta_j : j \in J\}$ ，分隔向量為 $c = \{c(\alpha_i, \beta_j) : i \in I, j \in J\}$ ，介於行為者 α_i 和機會 β_j ， $\{\alpha_{in} : i \in I\}, \{\beta_{jn} : j \in J\}, n = 1, \dots, N$ ，觀察的交互作用(observed interactions) $N(\alpha_i, \beta_j)$ ，作用頻次(interaction frequency) $N(\alpha_i, \beta_j)/N$ ，以及估計介於 α_i 和 β_j 的機率為 $P_c(\alpha_i, \beta_j)$ 。以下為 Ashish Sen 和 Tony E. Smith(1995)根據假設，所提出的一般性質：

一般性:

R1.(恆正性)

$\forall c \in C, \forall s \in S$, 型態機率 $P_c(s)$ 恆正。

R2.(對稱性)

給定 $\forall c \in C$, 和一對 $s', s \in S$, 若 s', s 的相異僅有排序上的不同 , 則 $P_c(s) = P_c(s')$ 。

R3.(連續性)

給定 $\forall s \in S$, $\forall c \in C$, 和 $\forall \varepsilon > 0$, 必定 $\exists \delta > 0$, 使得在所有在 $c' \in C$, $\|c - c'\| < \delta$ 時 , $|P_c(s) - P_{c'}(s)| < \varepsilon$ 恆成立。

Ashish Sen 和 Tony E. Smith(1995)更進一步定義空間作用過程 (Spatial interaction process)如下:

定義 3.1 空間作用過程(Spatial interaction process)

在 S 中的機率函數集合 , $\mathbf{P} = \{P_c : c \in C\}$, 是為空間作用過程 , 若且為若 , \mathbf{P} 滿足一般性 **R1**、**R2**、**R3**。

3.2.2 阻抗函數

阻抗函數(deterrence function)是測度行為者針對變動區位 i 和區位 j 距離的意願，阻抗函數有多種型態，以下介紹最常見:

(1) 冪次阻抗函數(Power deterrence function)

$$F(c_{ij}) = (c_{ij})^{-\theta} \quad (3-6)$$

(2) 指數阻抗函數(Exponential deterrence function)

$$F(c_{ij}) = \exp[-\theta c_{ij}] \quad (3-7)$$

其中 θ 可解釋為成本敏感度參數(cost sensitivity parameter)，

Kulldorf(1955)研究首度發現指數型態的阻抗函數比冪次型態更加配適諸如遷移旅次這類型的資料；這種根據經驗所獲得的結果，在 Morrill 和 Pitts(1967)的實驗中獲得驗證。

3.2.3 定義引力模式

Ashish Sen 和 Tony E. Smith(1995)將引力模式的定義如下:

定義 3.2: (一般引力模式)

$P = \{P_c : c \in C\}$ 被稱為引力模式，若且為若， $\forall c \in C$ ， $\exists A_c, B_c, F_c > 0$ ，使得所有起迄對， $ij \in I \times J$

$$E_c(N_{ij}) = A_c(i)B_c(j)F_c(c_{ij}) \quad (3-8)$$

定義 3.3: (指數型態引力模式)

$P = \{P_c : c \in C\}$ 被稱為指數型引力模式，若且為若， $\forall c \in C$ ， $\exists A_c, B_c > 0$ 和存在成本敏感度向量 $\theta_c \in R^K$ ，使得所有起迄對， $ij \in I \times J$

$$E_c(N_{ij}) = A_c(i)B_c(j)\exp[-\theta_c^t c_{ij}] \quad (3-9)$$

3.2.4 模式的原則

已知行為者總和 A ，機會總和 B ，分隔測度(separation measures), $\{c^k : k \in K\}$ ，定義空間總合組合(spatial aggregation scheme)具有一對有限分割的 $\{\alpha_i : i \in I\}$ 和 $\{\beta_j : j \in J\}$ ，總合函數 $\{\Phi^k : k \in K\}$ ，指派到每對 A_i 和 B_j ，則分隔量表示為 $c_{ij} = (c_{ij}^k : i \in I, j \in J)$ ，並以數學式表示：

$$c_{ij}^k = \Phi^k [c^k(\alpha, \beta) : \alpha \in A_i, \beta \in B_j] \quad , k \in K \quad (3-10)$$

則定義 c_{ij}^k 的集合分隔量

$$c_{ij} = (c_{ij}^k : k \in K) \quad (3-11)$$

表示每對 $((\alpha_i, \beta_j) : i \in I, j \in J)$ 的成本總和，進而的，將所有對發生的成本總合分隔結構(separation configuration)表示如下：

$$c = (c_{ij} : ij \in I \times J) \quad (3-12)$$

旅次調查之目的在了解由某一區 i 到另一區 j 的交通量，調查之後通常將資料表示成二向列聯表的形式(two-way contingency table)。首先介紹各符號的意義：

- T_{ij} : 路網中起點 i 至迄點 j 的旅次需求量。
- T : 旅次需求量的總和。
- O_i : i 區產生的旅次量。
- D_j : j 區吸引的旅次量。

在交通上，一個 IXJ 的二向列聯資料型態如下：

表 3-1 旅次起迄分佈矩陣表

D_j O_i	1	\dots	j	\dots	J	T_i
1	T_{11}	\dots	T_{1j}	\dots	T_{1J}	O_1
$:$	$:$		$:$		$:$	$:$
l	T_{l1}		T_{lj}		T_{lJ}	O_l
$:$	$:$		$:$		$:$	$:$
I	T_{I1}	\dots	T_{Ij}	\dots	T_{IJ}	O_I
T_j	D_1	\dots	D_j	\dots	D_J	T

由 3-1, 在(i,j)上的觀察值為 T_{ij} , 其中

$$T_{i+} = \sum_j T_{ij} = O_i \text{ 為 } i \text{ 列之和,}$$

$$T_{+j} = \sum_i T_{ij} \text{ 為 } j \text{ 行之和,}$$

T 為總數和。

Wilson(1967)發現，引力模型可用最大熵法(entropy maximizing)導出，使得此模型更具推理基礎，且改稱此種模型為空間交互作用模型(spatial interaction model)。

給定:

$$\sum_j T_{ij} = O_i \tag{3-13}$$

$$\sum_i T_{ij} = D_j \tag{3-14}$$

利用觀察值 O_i 、 D_j 的資訊來估計參數 $A(i)$ 、 $B(j)$ 和 θ ，以模式

$$T_{ij} = A(i)B(j)\exp[-\theta c_{ij}] \quad (3-15)$$

推估旅次 T_{ij} 。

推估的方法為根據限制式，極大化起迄矩陣 $\{T_{ij}\}$ 的熵分配，根據限制式多寡，Wilson 將它分類為以下四種引力模式：

(1) Unconstrained

$$T_{ij} = A_i O_i f(c_{ij}) \quad (3-16)$$

(2) Production-constrained

$$T_{ij} = A_i O_i f(c_{ij}) \quad (3-17)$$

其中 $A_i = \left[\sum_j f(c_{ij}) \right]^{-1}$ 。

(3) Attraction-constrained

$$T_{ij} = B_j D_j f(c_{ij}) \quad (3-18)$$

其中 $B_j = \left[\sum_i f(c_{ij}) \right]^{-1}$

(4) Doubly-constrained

$$T_{ij} = A_i B_j O_i D_j f(c_{ij}) \quad (3-19)$$

其中 $A_i = \left[\sum_j B_j D_j f(c_{ij}) \right]^{-1}$ (3-20)

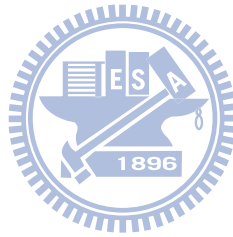
$B_j = \left[\sum_i A_i O_i f(c_{ij}) \right]^{-1}$ 以迭代方式求得 (3-21)

式(3-19)可改寫為 O_i 分派的雙限制式引力模式:

$$T_{ij} = O_i \frac{B_j D_j f(c_{ij})}{\sum_j B_j D_j f(c_{ij})}$$

式(3-19)可改寫為 D_j 分派的雙限制式引力模式:

$$T_{ij} = D_j \frac{A_i O_i f(c_{ij})}{\sum_i A_i O_i f(c_{ij})}$$



3.3 本研究模式

根據以上所述，我們將 I 視為 O_i ，即行為者為起點的產生量； J 視為 D_j ，意即將迄點的吸引量視為行為者欲達成的機會；再仔細探討 D_j 在整個起迄資料的關係， j 區位提供的經濟活動多寡，關係著 j 的吸引量 D_j ，簡言之，在 i 點有行為者欲到 j 點進行經濟活動，而 D_j 的交通量正可以類比 j 區所提供的經濟活動。

因上述關係，所以讓筆者感興趣於進階探討引力模式是否能經修正更加配適於交通起迄資料的推估。以下是關於引力模式的建構原則、假設以及性質：

3.3.1 模式建構的原則

給定起點群體(origin population), A 和迄點群體(destination population), B ，其中，存在一個相關的有限分隔屬性集合(separation attribute) $k \in K$ ，則測度分離(separation measures)集合為 $\{c^k : k \in K\}$ ；

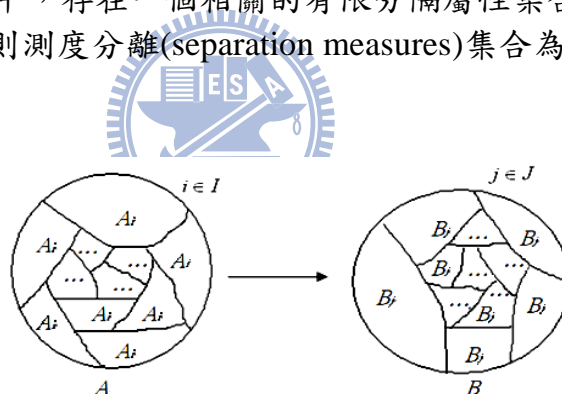


圖 3-1 子群體示意圖

[資料來源:本研究整理]

A_i 為 A 的子群體， B_j 為 B 的子群體，如上圖 3-1 所示， $i \in I$ ， $j \in J$ ，起點集合(origin set) I ，迄點集合(destination set) J ，定義空間總合系統(spatial aggregation scheme)為包含

A 和 B 的有限分割

$$\{A_i : i \in I\} \{B_j : j \in J\}$$

具有總合函數

$$\{\Phi^k : k \in K\} \quad (3-22)$$

指派到每對 A_i 和 B_j ，則分隔量為

$$c_{ij} = \{c_{ij}^k : i \in I, j \in J, k \in K\} \quad (3-23)$$

並以數學式表示：

$$c_{ij}^k = \Phi^k [c^k(\alpha, \beta) : \alpha \in A_i, \beta \in B_j] \quad , k \in K \quad (3-24)$$

上述，表示每對 $\{(\alpha_i, \beta_j) : i \in I, j \in J\}$ 的分隔集合， $c^k(\alpha, \beta)$ 表示介於 α, β 的 k 類分隔屬性。

進而的，將所有對發生的分隔總合分隔結構(separation configuration)表示如下：

$$c = \{c_{ij} : ij \in I \times J\} \quad (3-25)$$

再者， R 表示實數， R_+ 表示非負整數，若在 K 的非負分隔測度表為 K_+ ，則可能的分隔向量， V ，寫為卡氏積(cartesian product)型態

$$V = (R_+)^{K_+} \times (R)^{K-K_+} \quad (3-26)$$

進階的，假設相關的結構類別(configuration class), C

$$C = V^{I \times J} = \{c : c_{ij} \in V, ij \in I \times J\} \quad ij \in I \times J \quad (3-27)$$

並定義所有作用型態集合(出象空間,outcome space) 為 S 。

3.3.2 模式建構的基本假設

假設一：

假設所有子群體足夠大，以確保介於 A_i 和 B_j 的作用所產生的機率為正，即 $\forall s \in S, \forall c \in C, P_c(s) > 0$ 。

假設二:

假設所有 $\alpha \in A_i$ 對於某一個 β 所產生的分隔一致，且所有 $\beta \in B_j$ 對於某一個 α 亦所產生的分隔一致。

假設三:

假設 $\forall s \in S, \forall c \in C$, 使得 $P_c(s)$ 為一連續型函數。

3.3.3 模式的特性

本模式的一般性:

恆正性:

根據假設一，可得到機率恆正的性質；相對於到本研究的行為意義為在交通起迄資訊上結構性不為零的資料，必有起迄事件的發生，與實際情況一致，符合上述 R1. 恆正性。

對稱性:

根據假設二，可得到符合 R2. 對稱性的性質，了解到起迄產生的機率與時間為獨立，可以在下文經過完整定義後得到證明。

連續性:

根據假設一，解釋機率模型為一平滑曲線，使得針對模式作參數校估時，進行分割求解，仍能得到一個理想的解，亦符合上述 R3. 連續性。

根據 Sen & Smith(1955)所定義的空間作用過程，因為本模式滿足 R1、R2、R3，故機率函數集合

$$P = \{P_c : c \in C\}$$

是為一空間作用過程。

為了分析空間作用過程, $\mathbf{P} = \{P_c : c \in C\}$ 在 Ω 空間中的量測特性, 指定一數值函數 X 在 Ω 上, 在作用型態, w , 上可能發生的數值 x , 使得 $X(w) < x$; 令 $S(x) = \{s \in S : X(s) \leq x\}$, 則機率表示為

$$P_c(X \leq x) = \sum_{s \in S} P_c(s) \quad (3-28)$$

關聯的平均值(或是期望值)為

$$E_c(X) = \sum_{s \in S} X(s)P_c(s) \quad (3-29)$$

對本研究而言, 給定一個起迄對(origin-destination pair), 其相對應作用 (*ij-interactions*) 得作用旅次為 $T_{ij}(s)$, 對於一個非負整數, n_{ij} , 其機率為

$$P_c(n_{ij}) = P_c(T_{ij} = n_{ij}) = \sum_{s \in S} P_c(s) \quad (3-30)$$

則作用旅次為

$$E_c(T_{ij}) = \sum_{s \in S} T_{ij}(s)P_c(s) = \sum_{n_{ij}} n_{ij}P_c(n_{ij}) \quad (3-31)$$

總旅次記為

$$T(s) = \sum_{ij \in I \times J} T_{ij}(s) = T \quad (3-32)$$

若兩個作用型態, s 和 s' 的各別作用僅有發生順序上的差別, $\forall s, s' \in S$, 則

$$\begin{aligned} T_{ij}(s) &= T_{ij}(s') & ij \in I \times J \\ \Rightarrow P_c(s) &= P_c(s') \end{aligned} \quad (3-33)$$

故符合 **R2** 的對稱性。

若 $P_c(ij)$ 記為在 $ij \in I \times J (= S_1)$ 的型態機率(pattern probability), 則

$$P_c(S_1) = \sum_{ij \in I \times J} P_c(ij) \quad (3-34)$$

定義 3.4: (廣義一般型態引力模式)

$P = \{P_c : c \in C\}$ 被稱為引力模式，若且為若， $\forall c \in C, \exists A_c, B_c, F_c > 0$ ，使得起迄對 $(i, j), \forall ij \in I \times J$ ，

$$E_c(T_{ij}) = W_1(i)W_2(j)A_c^\delta(i)B_c^\gamma(j)F_c(c_{ij}) \quad (3-35)$$

其中 $W_1(i), W_2(j)$ 為校正因子， δ, γ 為模式參數。

定義 3.5: (廣義指數型態引力模式)

$P = \{P_c : c \in C\}$ 被稱為指數型引力模式，若且為若， $\forall c \in C, \exists A_c, B_c > 0$ ，且存在成本敏感度向量， $\theta_c \in R^K$ ，使得所有起迄對 $(i, j), \forall ij \in I \times J$ ，

$$E_c(T_{ij}) = W_1(i)W_2(j)A_c^\delta(i)B_c^\gamma(j)\exp[-\theta_c^t c_{ij}] \quad (3-36)$$

其中 $W_1(i), W_2(j)$ 為校正因子， δ, γ 為模式參數。

對於觀察該模式的行為意義而言，了解每個交互作用旅次對整個交通起迄資料的影響，因此，以機率型式表達，更具有重要的意義，作用機率(interaction probability)定義為

$$p_c(ij) = \frac{E_c(T_{ij})}{\sum_{gh \in I \times J} E_c(T_{gh})} = \frac{A_c(i)B_c(j)F_c(c_{ij})}{\sum_{gh \in I \times J} A_c(i)B_c(j)F_c(c_{gh})} \quad (3-37)$$

定義 3.3.3 的廣義指數型態引力模式寫為

$$p_c(ij) = \frac{E_c(T_{ij})}{\sum_{gh \in I \times J} E_c(T_{gh})} = \frac{W_1(i)W_2(j)A_c^\delta(i)B_c^\gamma(j)\exp[-\theta_c^t c_{ij}]}{\sum_{gh \in I \times J} W_1(i)W_2(j)A_c^\delta(i)B_c^\gamma(j)\exp[-\theta_c^t c_{ij}]} \quad (3-38)$$

根據定義 3.3.3 的廣義指數型態引力模式，當

case 1: $W_1(i) = 1$ ，且 $\delta = 0, \gamma = 1$ ，則

$$E_c(T_{ij}) = W_2(j)B_c(j)\exp[-\theta_c^t c_{ij}]$$

機率型態改寫為

$$p_c(ij) = \frac{E_c(T_{ij})}{\sum_{gh \in I \times J} E_c(T_{gh})} = \frac{W_2(j)B_c(j)\exp[-\theta_c^t c_{ij}]}{\sum_{gh \in I \times J} W_2(h)B_c(h)\exp[-\theta_c^t c_{ij}]}$$

則各起迄對 $(i, j), \forall ij \in I \times J$, 旅次 T_{ij} 為

$$\begin{aligned} T_{ij} &= A_c(i)p_c(ij) \\ &= A_c(i) \frac{E_c(T_{ij})}{\sum_{h \in J} E_c(T_{gh})} = A_c(i) \frac{W_2(j)B_c(j)\exp[-\theta_c^t c_{ij}]}{\sum_{h \in J} W_2(h)B_c(h)\exp[-\theta_c^t c_{ij}]} \\ &= W_1(i)W_2(j)A_c(i)B_c(j)\exp[-\theta_c^t c_{ij}] \end{aligned} \quad (3-39)$$

即為 3.2.4 節中，Wilson 所定義 Doubly constrained 的 O_i 分派引力模式；

case 2: $W_2(j) = 1$, 且 $\delta = 1, \gamma = 0$ ，則

$$E_c(T_{ij}) = W_1(i)A_c(i)\exp[-\theta_c^t c_{ij}]$$

機率型態改寫為

$$p_c(ij) = \frac{E_c(T_{ij})}{\sum_{gh \in I \times J} E_c(T_{gh})} = \frac{W_1(i)A_c(i)\exp[-\theta_c^t c_{ij}]}{\sum_{g \in I} W_1(i)A_c(i)\exp[-\theta_c^t c_{ij}]}$$

則各起迄對 $(i, j), \forall ij \in I \times J$, 旅次 T_{ij} 為

$$\begin{aligned} T_{ij} &= B_c(j)p_c(ij) \\ &= B_c(j) \frac{E_c(T_{ij})}{\sum_{g \in I} E_c(T_{gh})} = B_c(j) \frac{W_1(i)A_c(i)\exp[-\theta_c^t c_{ij}]}{\sum_{g \in I} W_1(i)A_c(i)\exp[-\theta_c^t c_{ij}]} \\ &= W_1(i)W_2(j)A_c(i)B_c(j)\exp[-\theta_c^t c_{ij}] \end{aligned} \quad (3-40)$$

即為 3.2.4 節中，Wilson 所定義 Doubly constrained 的 D_j 分派引力模式。

3.3.4 模式推導

根據上述的模式原則與假設，在本節，以本研究所整理的國道一號

收費站交通量，根據所得的資訊及限制式，利用極大熵法來推導上述所定義的廣義指數型態引力模式。同時，由 3.2.2 的抗阻函數得到，指數型態的抗阻函數針對起迄類型的資料具有較佳的配適能力，故本模式選抗阻函數為指數型態。

模式推導，主要在推導在相關假設、限制式之下，校正因子所隱含的變數，以利在第五章模式驗證時，將模式作實際應用。

3.3.4.1 模式(一)推導

以下為模式推導：

根據定義 3.5 以及 Wilson 所提出的雙限制式引力模式

$$E_c(T_{ij}) = W_1(i)W_2(j)A_c^\delta(i)B_c^\gamma(j)\exp[-\theta_c c_{ij}]$$

當 $W_1(i) = 1$ 且 $\delta = 0$ ，則

$$E_c(T_{ij}) = W_2(j)B_c^\gamma(j)\exp[-\theta_c c_{ij}] \quad (3-41)$$

機率型態改寫為

$$p_c(ij) = \frac{E_c(T_{ij})}{\sum_{gh \in I \times J} E_c(T_{gh})} = \frac{W_2(j)B_c^\gamma(j)\exp[-\theta_c c_{ij}]}{\sum_{g \in I} W_2(j)B_c^\gamma(j)\exp[-\theta_c c_{ij}]}$$

則各起迄對 $(i, j), \forall ij \in I \times J$ ，旅次 T_{ij} 為

$$\begin{aligned} T_{ij} &= A_c(i)p_c(ij) \\ &= A_c(i) \frac{E_c(T_{ij})}{\sum_{g \in I} E_c(T_{gh})} = A_c(i) \frac{W_2(j)B_c^\gamma(j)\exp[-\theta_c c_{ij}]}{\sum_{g \in I} W_2(j)B_c^\gamma(j)\exp[-\theta_c c_{ij}]} \end{aligned}$$

$$\text{令 } W_1'(i) = \left[\sum_{g \in I} W_2(j)B_c^\gamma(j)\exp[-\theta_c c_{ij}] \right]^{-1}$$

$$\text{則 } T_{ij} = W_1'(i)W_2(j)A_c(i)B_c^\gamma(j)\exp[-\theta_c c_{ij}]$$

根據限制式，因此

$$\begin{aligned} \sum_i T_{ij} &= \sum_i W_1'(i)W_2(j)A_c(i)B_c(j)\exp[-\theta_c c_{ij}] = B_c(j) \\ \Rightarrow W_2(j)B_c^r(j) \sum_i W_1'(i)A_c(i)\exp[-\theta_c c_{ij}] &= B_c(j) \\ \Rightarrow W_2(j) &= B_c(j) \left[B_c^r(j) \sum_i W_1'(i)A_c(i)\exp[-\theta_c c_{ij}] \right]^{-1} \end{aligned}$$

模式(一):

引力模式

$$T_{ij} = W_1'(i)W_2(j)A_c(i)B_c^r(j)\exp[-\theta_c c_{ij}]$$

其中

$$\begin{aligned} W_1'(i) &= \left[\sum_{g \in I} W_2(j)B_c^r(j)\exp[-\theta_c c_{ij}] \right]^{-1} \\ W_2(j) &= B_c(j) \left[B_c^r(j) \sum_i W_1'(i)A_c(i)\exp[-\theta_c c_{ij}] \right]^{-1} \end{aligned}$$

模式(一)為指數型引力模式。

由於交通量的產生，源自於運輸需求，而運輸需求是一種衍生性需求，而兩起迄點之間的經濟活動方才是交通量產生的泉源；因此，模式(一)的意義在於給予 O_i, D_j 一組敏感度參數 δ, γ ，使之彌補古典引力模式所無法解釋的母體實際現象。

3.3.4.2 模式(二)推導

已知當求解最佳化問題時，給予的限制式越多，越能解出符合所求的最佳解；在本研究中，所謂的最佳解，即是校估出一組最佳參數的引力模式，再利用樣本資料推估母體。

因此，在模式(二)中，希望加入一組路段限制式，使之推估出的參數，更能解釋母體現象；以下是關於線性路網(高速公路主線上收費站)路段資訊:(附錄、一)

$$\sum_k f_k = 2 \sum_i \sum_j T_{ij} - s$$

$$\text{其中 } s = \sum_m T_{m,m+1} + \sum_n T_{n+1,n}$$

f : 線性路網路段流量

將 doubly-constrained 加上路段限制式

$$\sum_k f_k = 2 \sum_i \sum_j T_{ij} - s$$

$$\text{其中 } s = \sum_m T_{m,m+1} + \sum_n T_{n+1,n}$$

推導過程如下:

目標式: 極大化起迄矩陣 $\{T_{ij}\}$ 的熵分配

$$W(\{T_{ij}\}) = \frac{T!}{\prod_{ij} T_{ij}!}$$

T 為總起迄量

故可能產生的情況是 w ，且其數量超過 T_{ij}

$$w = \sum W(\{T_{ij}\})$$

限制式:

$$\sum_j T_{ij} = A_c(i)$$

$$\sum_i T_{ij} = B_c(j)$$

$$\sum_i \sum_j T_{ij} C_{ij} = C$$

$$\sum_k f_k = 2 \sum_i \sum_j T_{ij} - s$$

以 Lagrangian multiplier 求解目標式:

$$\begin{aligned} \mathcal{L} = \ln W + \sum_i \lambda_i (A_c(i) - \sum_j T_{ij}) + \sum_j \lambda_j (B_c(j) - \sum_i T_{ij}) + \beta (C - \sum_i \sum_j T_{ij} C_{ij}) \\ + \gamma (\sum_k f_k - 2 \sum_i \sum_j T_{ij} - s) \end{aligned}$$

$$s = \sum_m T_{m,m+1} + \sum_n T_{n+1,n}$$

$\lambda_i, \lambda_j, \beta, \gamma$ 為 Lagrangian multipliers

因求解 $\ln W$ 比 W 有效率，故利用 Stirling 近似

$$\ln N! \approx N \ln N - N$$

則

$$\frac{\partial \ln N!}{\partial N} = \ln N$$

故

$$\text{令 } \frac{\partial \mathcal{L}}{\partial T_{ij}} = -\ln T_{ij} - \lambda_i - \lambda_j - \beta C_{ij} - 2\gamma = 0$$

$$\ln T_{ij} = -\lambda_i - \lambda_j - \beta C_{ij} - 2\gamma$$

$$T_{ij} = e^{-\lambda_i - \lambda_j - \beta C_{ij} - 2\gamma}$$

根據限制式

$$\sum_j T_{ij} = \sum_j e^{-\lambda_i - \lambda_j - \beta C_{ij} - 2\gamma} = A_c(i)$$

$$\text{令 } W_1(i) = \frac{e^{-\lambda_i}}{A_c(i)} = e^{2\gamma} \cdot \left[\sum_j e^{-\lambda_j - \beta C_{ij}} \right]^{-1} \quad (3-42)$$

$$\sum_i T_{ij} = \sum_i e^{-\lambda_i - \lambda_j - \beta C_{ij} - 2\gamma} = B_c(j)$$

$$\text{令 } W_2(j) = \frac{e^{-\lambda_j}}{B_c(j)} = e^{2\gamma} \cdot \left[\sum_i e^{-\lambda_i - \beta C_{ij}} \right]^{-1} \quad (3-43)$$

$$\sum_i \sum_j T_{ij} = \sum_i \sum_j e^{-\lambda_i - \lambda_j - \beta C_{ij} - 2\gamma} = \frac{1}{2} \left(\sum_k f_k - s \right)$$

$$\text{令 } L = e^{2\gamma} = \frac{2 \sum_i \sum_j W_1(i) W_2(j) A_c(i) B_c(j) f(C_{ij})}{\left(\sum_k f_k - s \right)} \quad (3-44)$$

則

$$\begin{aligned}
 T_{ij} &= e^{-\lambda_i - \lambda_j - \beta C_{ij} - 2\gamma} \\
 &= \frac{W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})}{\sum_i \sum_j W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})} \cdot \frac{1}{2} \left(\sum_k f_k - s \right) \\
 &= T \cdot G \cdot \frac{W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})}{\sum_i \sum_j W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})}
 \end{aligned}$$

$$\text{其中 } W_1(i) = e^{2\gamma} \cdot \left[\sum_j e^{-\lambda_j - \beta C_{ij}} \right]^{-1} = L \left[\sum_j W_2(j)B_c(j)f(C_{ij}) \right]^{-1}$$

$$W_2(j) = e^{2\gamma} \cdot \left[\sum_i e^{-\lambda_i - \beta C_{ij}} \right]^{-1} = L \left[\sum_i W_1(i)A_c(i)f(C_{ij}) \right]^{-1}$$

$$L = e^{2\gamma}$$

$f(c_{ij}) = \text{xp}[-\theta_c c_{ij}]$ 為阻抗函數。

$$\text{則 } T_{ij} = T \cdot \frac{2}{L(\sum_k f_k - s)} \cdot W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})$$

$$\text{令 } \frac{2}{L(\sum_k f_k - s)} = G$$

$$\text{故 } T_{ij} = T \cdot G \cdot W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})$$

(3-45)

根據定義 $E(T_{ij}) = W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})$

可知所推導的模式是引力模式，且為指數型引力模式。

模式(二):

$$T_{ij} = T \cdot G \cdot W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})$$

$$\text{其中 } G = \frac{2}{L(\sum_k f_k - s)}$$

$$W_1(i) = e^{2\gamma} \cdot \left[\sum_j e^{-\lambda_j - \beta C_{ij}} \right]^{-1} = L \left[\sum_j W_2(j)B_c(j)f(C_{ij}) \right]^{-1}$$

$$W_2(j) = e^{2\gamma} \cdot \left[\sum_i e^{-\lambda_i - \beta C_{ij}} \right]^{-1} = L \left[\sum_i W_1(i)A_c(i)f(C_{ij}) \right]^{-1}$$

$$L = e^{2\gamma}$$

$f(c_{ij}) = \text{xp}[-\theta_c c_{ij}]$ 為阻抗函數。

模式(二)亦為指數型引力模式，需校估參 $G, W_1(i), W_2(j)$ 。

第四章 抽樣分析與資料整理

4.1 抽樣方法

抽樣調查(sample surveys)是社會科學常用的一種研究方法，一般適用於群體、社會、地區，以至整個國家的研究。個別調查研究都有固定的研究範圍；研究人員利用資料蒐集工具，向範圍內的研究對象收集資料。但可能全數進行調查，所以在客觀的資源條件下，抽取適當的樣本，希望能從中獲得可靠的資料，以代表母體的實際情況，此乃抽樣調查的功能。

抽樣調查的方法已經遍佈於各個領域，對於旅次起迄量的調查者而言，日益增加的交通量，使得全面普查更加艱鉅，唯有靠著以機率為基礎的抽樣方法，才能改善調查的困難。因此，以下就針對旅次交通量的特性，提出一個可行的抽樣方法，以進行實際資料問題的分析。

4.1.1 系統抽樣

系統抽樣(Systematic sampling)，其抽樣方法是按母體中各元素之連續序列，有系統地每經一定間隔抽取一個元素作為調查對象的方法。設間距是 k ，若用 $i=1,2,3,\dots,k$ 分別代表由任某 i 值為起點用系統抽樣抽到的樣本，若我們將這些 k 個樣本視為 k 個集群，則母體即由這 k 個集群集成，因此系統抽樣即是由此 k 集群抽出一群為樣本(林進田,民 82)。

優點:

(1) 抽樣手續簡便易行，省掉編製底冊及抽取號碼的手續。母體之大小可以不必事先知道。

(2) 若母體元素之順序不呈某種週期序列，則系統抽樣的樣本

通常有很好的代表性，所做的估計精確性高。

使用限制:

- (1) 因系統抽樣是集群抽樣抽出一群的情況，故無法計算估計的誤差，只能由抽樣架構知道此估計的好壞。
- (2) 若母體元素之順序呈某種週期序列，則系統抽樣間隔與週期一致時或呈倍數關係，則所得樣本之代表性將受到質疑。

適用時機:

- (1) 當母體還沒形成就要抽樣時。
- (2) 當母體元素是不可數時。
- (3) 當母體很大，不呈周期性排列時，可採本法較為簡單。

雖然系統抽樣有其使用限制，但由於母體個數無法事先預知、元素之順序組成不呈週期序列時，系統抽樣是一很好的調查方法，而本研究之旅次起迄分佈矩陣之調查方法，正滿足上述之條件，因此本研究將以系統抽樣之調查方法，並進一步探討問題。

4.2 檢定方法

在往後的研究當中，需要藉由檢定方法來檢視(1) 不同抽樣比例的樣本與觀察資料的分配是否相同；(2) 所推估的起迄表與觀察資料的分配是否相同；此處我們採用 Kolmogorov-Smirnov 檢定法，簡稱 K-S 檢定法。此法可作適合度檢定，亦可用於特定階層之樣本與母體比例的比較或兩樣本之間的分配比較。

K-S 檢定法與卡方檢定法皆可作適合度檢定，以下針對兩者作比較，其差異為：

- (1) K-S 檢定法以機率或相對次數為依據，進行檢定工作，可檢定效率受樣本大小影響。
- (2) 不要求理論次數需大於等於 2 或 5，故不必將分組資料進行合併，可保持原有資料的面貌。
- (3) 在小樣本情形下，卡方檢定法不能使用，K-S 檢定法仍然適用。
- (4) K-S 檢定法可用於檢定某一階層的特定性質結構是否與母體一致。
- (5) K-S 檢定法的使用必須要求母體已知，而卡方檢定法可以統计量估計母體，而進行檢定。

根據本研究的資料特性，故選擇 K-S 檢定法。

4.3 資料蒐集

本研究資料來自遠通電收 2006 年至 2007 年高速公路所蒐集的資料，本研究範圍為國道一號中山高速公路，故儘擷取國道一號所偵測得的資料。原始資料如圖 4-1，檔案格式為.txt 文字檔，每天每五分鐘輸出一筆文字檔，一天共 288 個文字檔。

28307045,01002,07,2007-06-30T23:48:44,31,N

28307092,01002,07,2007-06-30T23:49:06,31,N

28307122,01002,07,2007-06-30T23:50:10,31,N

圖 4-1 原始資料型態

[資料來源:遠通電收]

檔案內容包含由左至右以逗號相隔分別為:車輛流水號、收費站代號(不分北上南下如表 4-1)、車道、偵測時間、車種(如表 4-2)、方向(如表 4-3)。

表 4-1 收費站代碼

收費站名稱	現有代碼
汐止收費站	04002
泰山收費站	14003
楊梅收費站	22002
造橋收費站	27002
月眉收費站	14004
后里收費站	30002
員林收費站	40002
斗南收費站	44002
新營收費站	48002
新市收費站	52003
岡山收費站	57002
註:其中汐止收費站南下不收費;月眉收費站北上出口以及南下入口不收費。	

[資料來源:本研究整理]

表 4-2 車種代碼

代碼	車種
31	小客車
32	小貨車
41	大客車
42	大貨車
5	聯結車
0	其他

[資料來源:本研究整理]

表 4-4 方向

代號	方向
N	北上
S	南下
E	東
W	西

[資料來源:本研究整理]

由於本研究著重在交通車輛起迄資料推估，且車輛起迄表不分車種，故忽略車種。

4.2.1 資料整理

本研究資料為民國 96 年 7 月 6 日高速公路遠通電信 ETC 資料，由於月眉收費站北上出口以及南下入口不收費，且其收費站設於交流道上，影響起迄表推估，故將月眉站移除。其中因汐止收費站南下不收費，導致北上至汐止收費站偵測的量為 0；又月眉收費站北上出口以及

南下入口不收費，故北上至月眉收費站和自月眉收費站南下的量亦為 0，以上乃是結構上為 0 的不完整資料，亦即母體本身資料即為 0，此種不完整資料是事先知道的，非本文研究重點，故不予以處理(如表 4-5 粗體字)。

車輛起迄表起迄對定義為自第一站(汐止站) 96 年 7 月 6 日 0 時起持續整理 28 小時資料，意即整理跨夜資料，各站根據時間遞移抽樣 28 小時，如下圖所示：

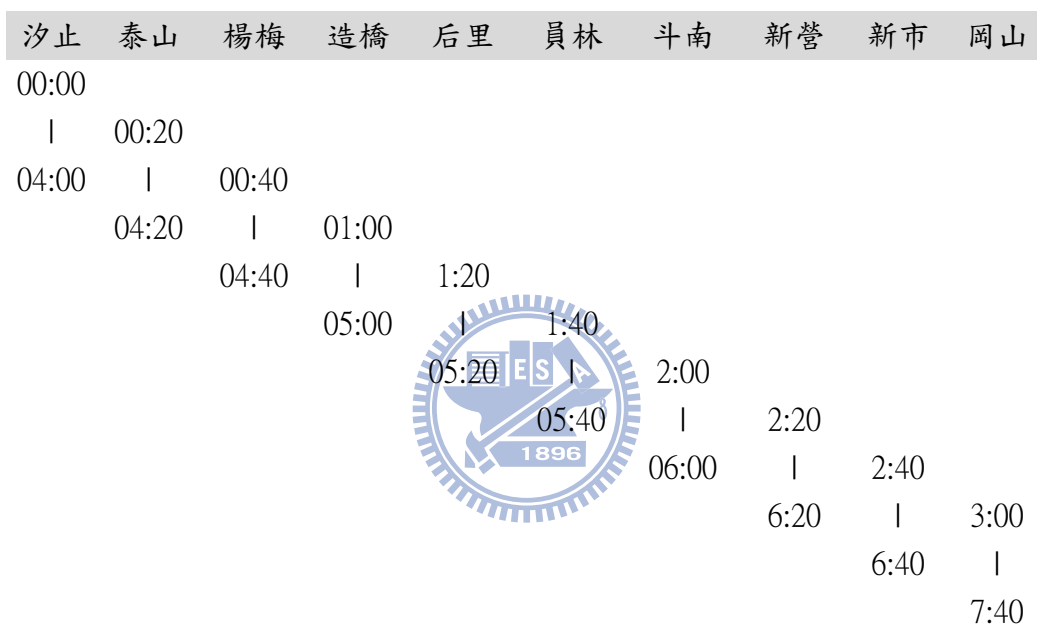


圖 4-2 時間遞移抽樣 28 小時

[資料來源:本研究整理]

交通車輛起迄表整理如表 4-5:

表 4-5 國道一號 交通車輛起迄表

	汐止	泰山	楊梅	造橋	后里	員林	斗南	新營	新市	岡山	列總和
汐止	0	874	137	31	64	23	7	10	1	2	1149
泰山	0	0	2468	271	1270	263	207	92	276	323	5170
楊梅	0	2291	0	536	1450	348	150	174	183	353	5485
造橋	0	265	438	0	2149	308	121	169	213	236	3899
后里	0	2434	2019	3325	0	95	92	108	106	194	8373
員林	0	92	119	135	59	0	947	717	882	1483	4434
斗南	0	48	51	60	25	518	0	389	450	924	2465
新營	0	36	46	55	23	295	1539	0	637	780	3411
新市	0	33	56	69	25	314	1729	527	0	1208	3961
岡山	0	27	53	67	25	268	3109	753	1401	0	5703
行總和	0	6100	5387	4549	5090	2432	7901	2939	4149	5503	44050

[資料來源:本研究整理]

此完整起迄表在第五章作為實例驗證之用。

4.2.2 抽樣比例與母體型態分析

由於普查耗費時間、人力，且常礙於經費限制，故需藉由抽樣所得有效的樣本來代表母體本身。本研究對母體進行抽樣，每 2 輛抽一輛、每 10 輛抽一輛、每 20 輛抽一輛、每 30 輛抽一輛。每 2 輛抽一輛針對各個收費站，抽取經過該收費站的第 1 輛、第 3 輛、第 5 輛.....，根據文獻可知抽取點的不同並不影響分配型態，亦即調查者於實際系統抽樣調查時，不論以何點抽起，(每 2 輛的第 2 輛)，其所得之旅次量(量)與其型態(質)是不會影響所得結果，(李霞,民 85)。

以下對母體進行不同抽樣比例抽樣：

表 4-6 每 2 輛車抽取第一輛車的起迄表

	汐止	泰山	楊梅	造橋	后里	員林	斗南	新營	新市	岡山	列總和
汐止	0	228	41	13	19	8	5	5	1	1	321
泰山	0	0	712	259	368	111	97	79	104	96	1826
楊梅	0	774	0	461	614	194	117	148	143	117	2568
造橋	0	402	535	0	925	198	136	126	153	105	2580
后里	0	600	827	1259	0	129	101	116	127	113	3272
員林	0	40	60	61	58	0	516	465	433	423	2056
斗南	0	17	39	41	37	231	0	399	481	452	1697
新營	0	16	29	31	31	141	860	0	499	440	2047
新市	0	16	21	40	27	140	881	639	0	485	2249
岡山	0	6	16	15	13	78	819	597	689	0	2233
行總和	0	2099	2280	2180	2092	1230	3532	2574	2630	2232	20849

[資料來源:本研究整理]



表 4-7 每 10 輛車抽取第一輛車的起迄表

	汐止	泰山	楊梅	造橋	后里	員林	斗南	新營	新市	岡山	列總和
汐止	0	11	1	0	0	0	0	1	0	0	13
泰山	0	0	37	24	20	11	11	10	5	3	121
楊梅	0	61	0	50	32	21	9	12	6	8	199
造橋	0	20	53	0	78	21	15	15	21	11	234
后里	0	28	39	68	0	24	14	20	10	14	217
員林	0	3	4	5	3	0	46	36	43	18	158
斗南	0	4	0	3	4	15	0	59	49	34	168
新營	0	2	3	7	3	8	64	0	49	36	172
新市	0	0	0	2	3	11	52	67	0	42	177
岡山	0	0	0	0	1	5	25	31	47	0	109
行總和	0	129	137	159	144	116	236	251	230	166	1568

[資料來源:本研究整理]

表 4-8 每 20 輛車抽取第一輛車的起迄表

	汐止	泰山	楊梅	造橋	后里	員林	斗南	新營	新市	岡山	列總和
汐止	0	4	0	0	0	0	0	1	0	0	5
泰山	0	0	9	8	6	5	2	4	1	1	36
楊梅	0	13	0	13	8	3	3	2	2	1	45
造橋	0	6	15	0	29	11	3	1	7	1	73
后里	0	10	9	24	0	6	3	10	4	8	74
員林	0	1	0	1	1	0	8	11	15	6	43
斗南	0	0	0	2	1	2	0	20	16	16	57
新營	0	0	0	3	0	4	19	0	17	11	54
新市	0	0	0	1	0	4	16	20	0	14	55
岡山	0	0	0	0	0	0	7	9	7	0	23
行總和	0	34	33	52	45	35	61	78	69	58	465

[資料來源:本研究整理]



表 4-9 每 30 輛車抽取第一輛車的起迄表

	汐止	泰山	楊梅	造橋	后里	員林	斗南	新營	新市	岡山	列總和
汐止	0	1	0	0	0	0	0	0	0	0	1
泰山	0	0	7	2	2	1	1	2	0	0	15
楊梅	0	14	0	6	6	2	3	1	1	0	33
造橋	0	4	6	0	12	5	3	0	1	0	31
后里	0	4	4	6	0	4	5	0	2	1	26
員林	0	0	1	1	0	0	11	4	6	1	24
斗南	0	0	0	2	0	3	0	14	6	4	29
新營	0	0	0	1	1	0	10	0	7	3	22
新市	0	0	0	0	1	1	9	9	0	7	27
岡山	0	0	0	0	1	1	1	2	5	0	10
行總和	0	23	18	18	23	17	43	32	28	16	218

[資料來源:本研究整理]

以上兩組樣本資料作為第五章模式驗證時用。

第五章 模式推估與驗證

本研究所提之模式已於第三章描述，本章延續第四章所整理之國道一號實際資料，以第三章所提之模式進行旅次分配推估與抽樣所產生不完整資料處理，進而將處理後之資料適當的放大與母體進行型態的檢定。

5.1 樣本與母體型態分析

根據4.2節的K-S檢定法以下針對各抽樣比例的樣本與母體型態作分析，即針對各站抽樣後，再進行配對後的樣本起迄型態與母體型態的關係，若型態一致，再進行起迄資料推估，檢定結果如下：

H_0 :母體與各抽樣比例的樣本型態相同

H_1 :母體與各抽樣比例的樣本型態不盡相同

$\alpha = 0.05$

抽樣比例	個數 (母體,樣本)	檢定統計量 D	臨界值 $D_{\alpha/2}$	決策
每 2 筆抽一筆	81,81	0.077176	0.2137	接受
每 10 筆抽一筆	81,81	0.133206	0.2137	接受
每 20 筆抽二筆	81,81	0.161438	0.2137	接受
每 30 筆抽二筆	81,81	0.123738	0.2137	接受

由以上檢定可知，所抽取的各抽樣比例樣本與母體資料型態一致，故接著將資料作旅次分配推估以及第六章的不完整資料處理。

5.2 起迄資料推估流程

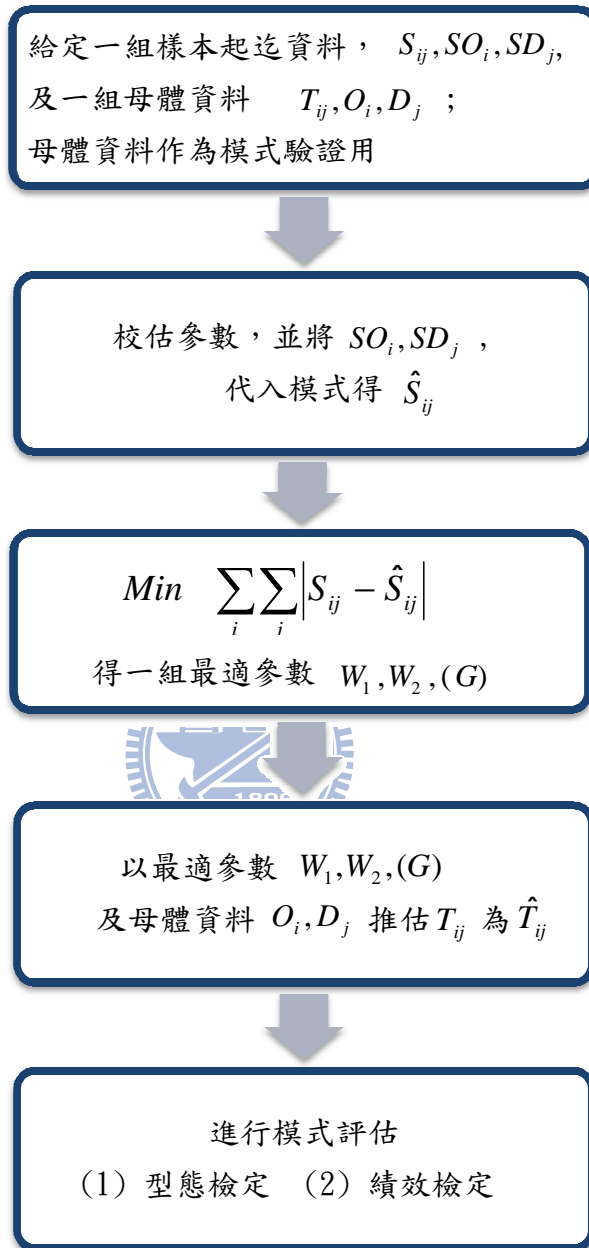


圖 5-1 起迄資料推估流程圖

[資料來源:本研究整理]

5.3 模式優劣程度指標

本節主要想了解經本研究提模式進行推估後，推估資料與母體資料的差異性。我們將抽象的指標予以量化，提出適用於本研究的評估指標。以下將提出本研究對於模式優劣評選的原則：

- (1) 一般比較、評估各種推估方法結果之精確性與差異，大多以

$$ERR = \sum_{ij} |T_{ij} - \hat{T}_{ij}| \quad \text{或} \quad ERR = \sum_{ij} (T_{ij} - \hat{T}_{ij})^2$$

為依據，所代表的意義大致為估計誤差的總和，其值越小表示推估的方法越理想。

- (2) 以下指標作為模式間優劣評估：


$$I = \frac{ERR_2 - ERR_1}{ERR_1} \times 100\%$$

I 值為正，表示模式 2 的 ERR 項比模式 1 大，即模式 1 較佳；

I 值為負，表示模式 2 的 ERR 項比模式 1 小，即模式 2 較佳。

5.4 模式推估與優劣程度測試

5.4.1 古典引力模式推估

以 Wilson 的雙限制式引力模式為例：

$$T_{ij} = A_i B_j O_i D_j f(c_{ij})$$

以下比較圖，為方便呈現，將 $T_{1,1}, T_{2,1}, \dots, T_{10,1}$ 分別以起迄對 $OD_1, OD_2, \dots, OD_{10}$ 表示，將 $T_{1,2}, T_{2,2}, \dots, T_{10,2}$ 分別以起迄對 $OD_{11}, OD_{12}, \dots, OD_{20}$ ，一直編號到 $T_{1,10}, T_{2,10}, \dots, T_{10,10}$ 分別以起迄對 $OD_{91}, OD_{92}, \dots, OD_{100}$ 。

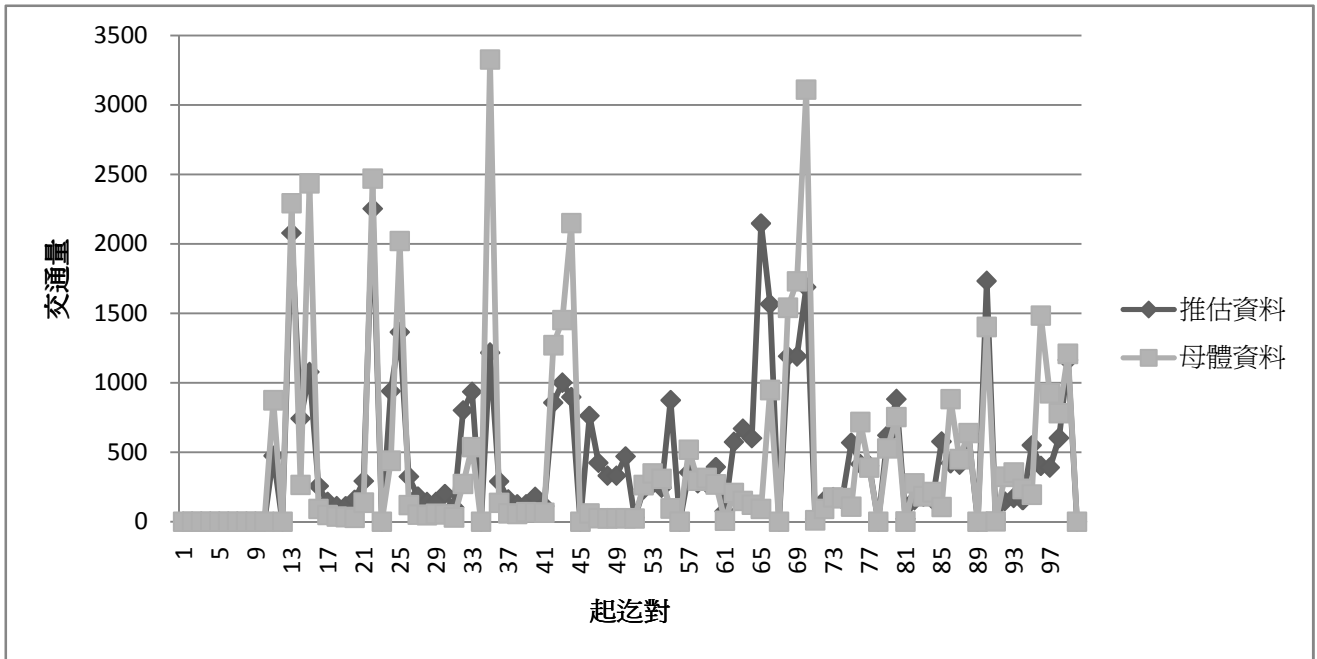


圖 5-2 (古典引力模式)母體與推估起迄資料比較圖 [資料來源:本研究整理]

模式指標：

$$ERR = \sum_{ij} |T_{ij} - \hat{T}_{ij}| = 26455$$

模式檢定：

H_0 :母體資料與推估資料型態相同

H_1 :母體資料與推估資料型態不盡相同

$$\alpha = 0.05$$

$$\text{統計量 } D_{\alpha/2} = 0.076909$$

$$\text{臨界值} = 0.2137$$

接受虛無假設，即母體資料與推估資料型態相同。

5.4.2 模式(一)

模式(一):

$$T_{ij} = W_1(i)W_2(j)A_c(i)B_c^y(j)\exp[-\theta_c^t c_{ij}]$$

分別對各抽樣分配所得的樣本做起迄資料推估，推估結果如下:

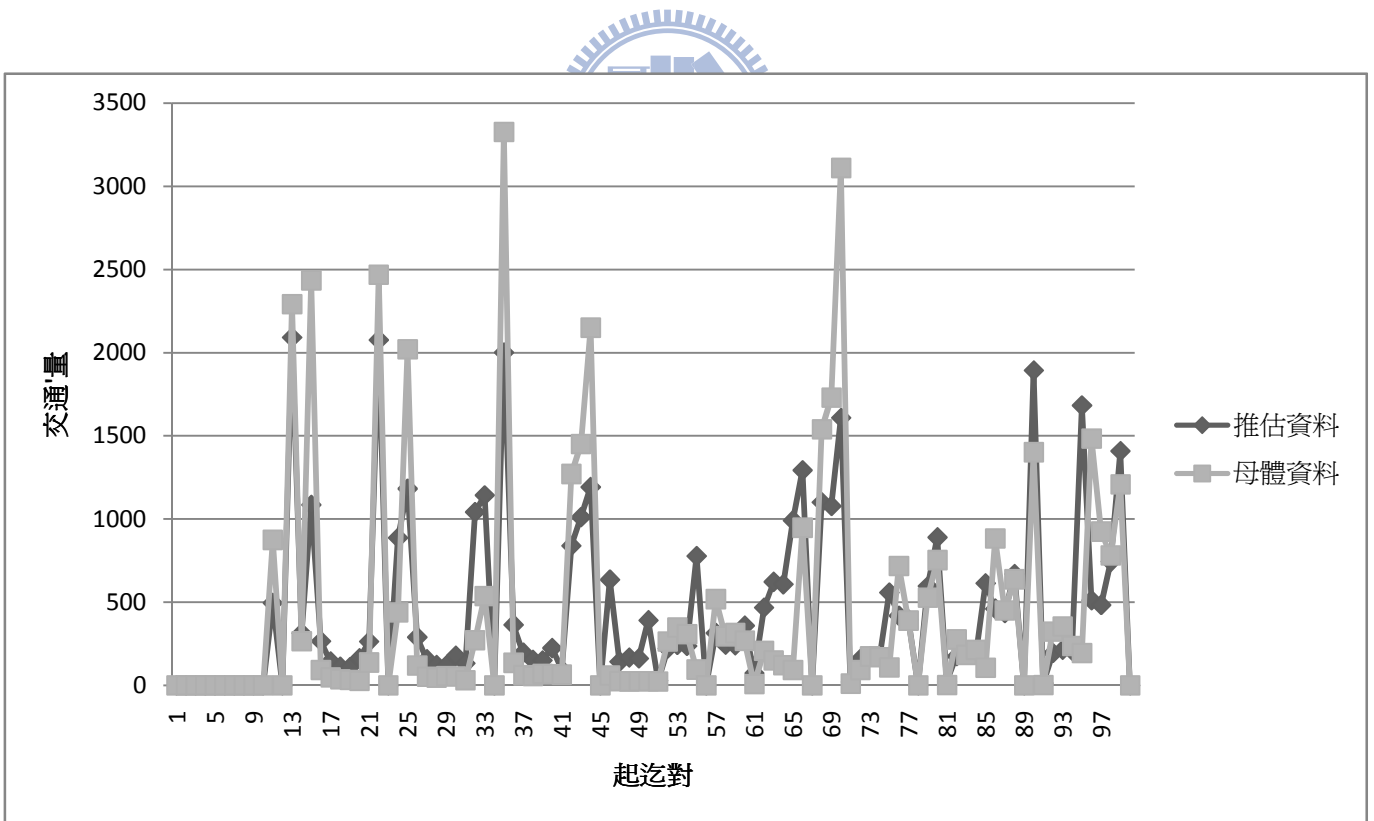


圖 5-3 模式(一)母體與推估起迄資料比較圖

[資料來源:本研究整理]

模式指標:

$$ERR = \sum_{ij} |T_{ij} - \hat{T}_{ij}| = 24128$$

模式檢定:

H_0 :母體資料與推估資料型態相同

H_1 :母體資料與推估資料型態不盡相同

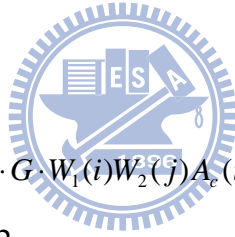
$$\alpha = 0.05$$

$$\text{檢定統計量 } D_{\alpha/2} = 0.05$$

$$\text{臨界值} = 0.2137$$

接受虛無假設，即母體資料與推估資料型態相同。

5.4.3 模式(二)



$$T_{ij} = T \cdot G \cdot W_1(i)W_2(j)A_c(i)B_c(j)f(C_{ij})$$

$$\text{其中 } G = \frac{2}{L(\sum_k f_k - s)}$$

$$W_1(i) = e^{2\gamma} \cdot \left[\sum_j e^{-\lambda_j - \beta C_{ij}} \right]^{-1} = L \left[\sum_j W_2(j)B_c(j)f(C_{ij}) \right]^{-1}$$

$$W_2(j) = e^{2\gamma} \cdot \left[\sum_i e^{-\lambda_i - \beta C_{ij}} \right]^{-1} = L \left[\sum_i W_1(i)A_c(i)f(C_{ij}) \right]^{-1}$$

$$L = e^{2\gamma}$$

$f(c_{ij}) = \exp[-\theta_c c_{ij}]$ 為阻抗函數。

模式參數可以遞迴方式求得。

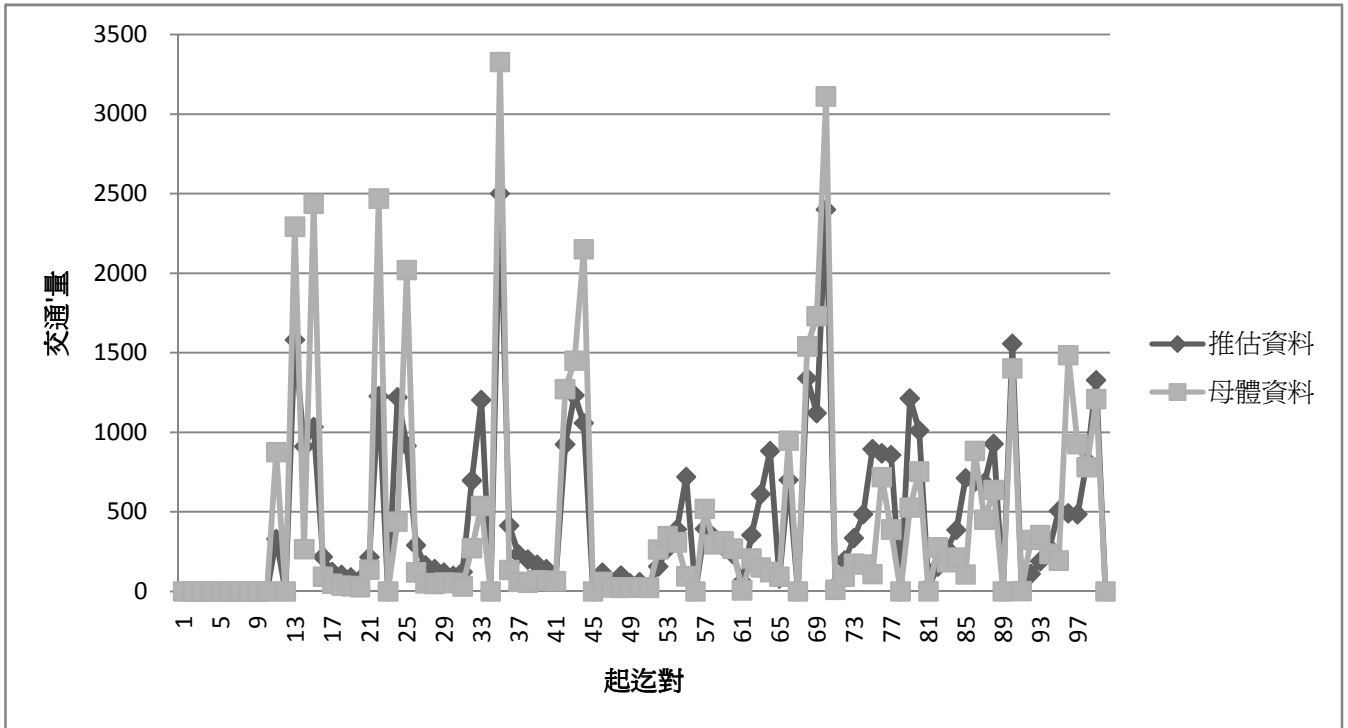


圖 5-4 模式(二)母體與推估起迄資料比較圖



[資料來源:本研究整理]

模式指標:

$$ERR = \sum_{ij} |T_{ij} - \hat{T}_{ij}| = 23899$$

模式檢定:

H_0 :母體資料與推估資料型態相同

H_1 :母體資料與推估資料型態不盡相同

$$\alpha = 0.05$$

$$\text{檢定統計量 } D_{\alpha/2} = 0.076162$$

$$\text{臨界值} = 0.2137$$

接受虛無假設，即母體資料與推估資料型態相同。

由古典引力模式與模式(一)的實驗結果可知，模式經修改給予 O_i, D_j 幕次 δ, γ 能夠更加配適原來的資料型態；藉由模式(一)與模式(二)可知，給予模式更多限制式，亦能夠獲得更加的推估結果。

上述實驗結果，可知模式經由修改之後，可獲得較佳的推估。

5.4.4 小抽樣比例樣本推估

在上一節中，分別以每 2 筆抽一筆的樣本資料作模式驗證，得到不錯的推估能力；在本節，以每 30 筆抽一筆的樣本資料作推估，以了解本研究模式，在小抽樣比例樣本之下，是否能得到好的推估能力。

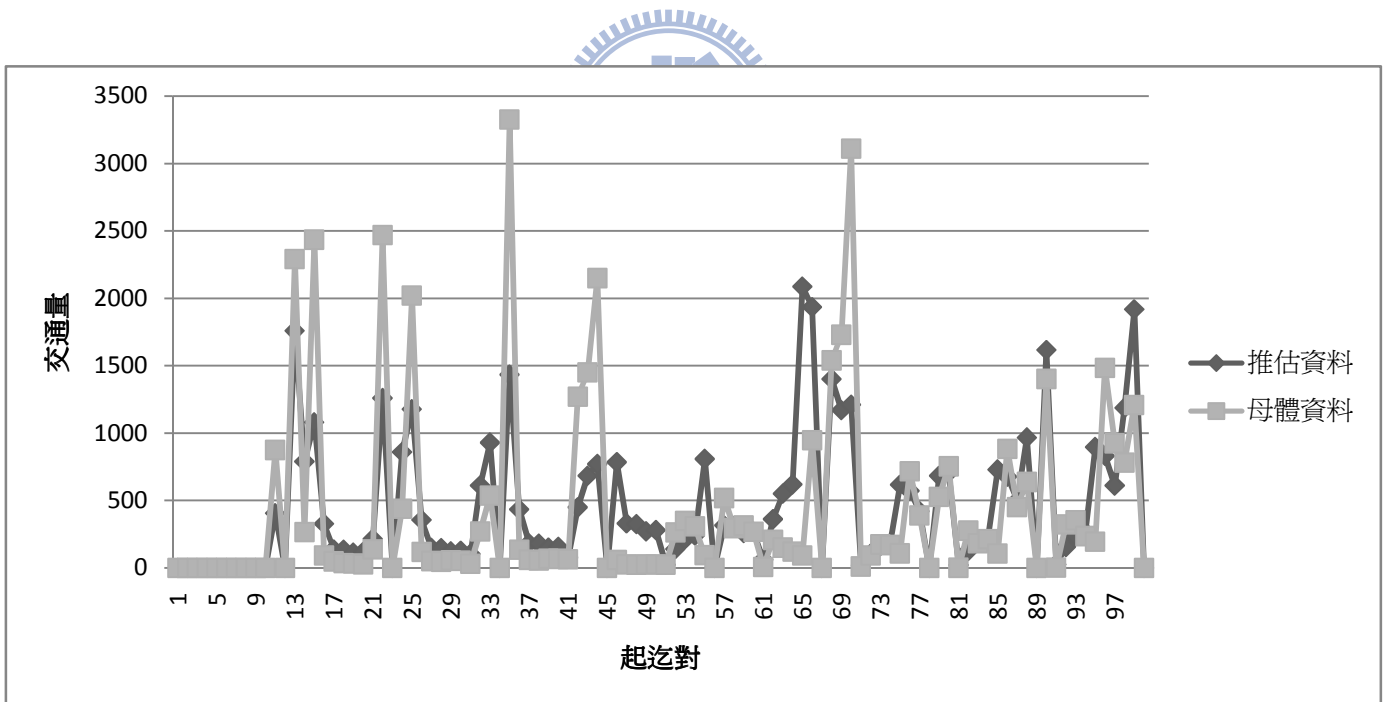


圖 5-5 古典引力:以抽樣比例小樣本推估資料與母體資料比較圖

[資料來源:本研究整理]

模式指標:

$$ERR = \sum_{ij} |T_{ij} - \hat{T}_{ij}| = 28490$$

模式檢定:

H_0 :母體資料與推估資料型態相同

H_1 :母體資料與推估資料型態不盡相同

$$\alpha = 0.05$$

$$\text{檢定統計量 } D_{\alpha/2} = 0.125529$$

$$\text{臨界值} = 0.2137$$

接受虛無假設，即母體資料與推估資料型態相同。

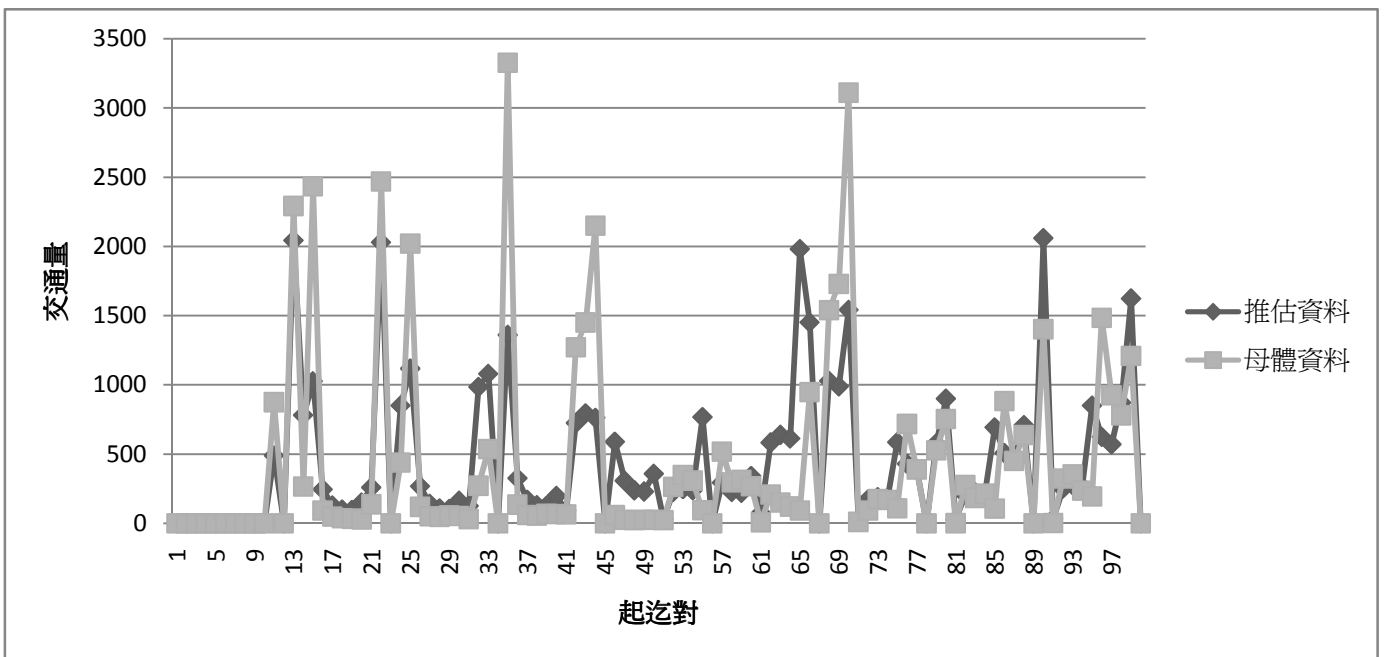


圖 5-6 模式(一):以抽樣比例小樣本推估資料與母體資料比較圖

[資料來源:本研究整理]

模式指標:

$$ERR = \sum_{ij} |T_{ij} - \hat{T}_{ij}| = 26923.1$$

模式檢定:

H₀:母體資料與推估資料型態相同

H₁:母體資料與推估資料型態不盡相同

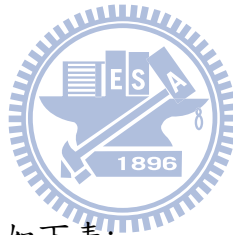
$$\alpha = 0.05$$

$$\text{檢定統計量 } D_{\alpha/2} = 0.091169$$

$$\text{臨界值} = 0.2137$$

接受虛無假設，即母體資料與推估資料型態相同。

由上述結果可知，當利用可獲得的小抽樣比例樣本時，仍能得到不錯的推估。



5.4.5 小結

將前述實驗結果整理如下表:

$$I = \frac{ERR_2 - ERR_1}{ERR_1} \times 100\%$$

其中 ERR_1 表古典引力模式推估資料與母體的變異

ERR_2 表本研究模式(一)(二)推估資料與母體的變異

表 5-1 模式間推估程度比較

	ERR	I	K-S 統計量	決策
古典引力模型	26455	-----	0.0769	接受
模式(一)	24128	-8.7961%	0.05	接受
模式(二)	23899	-9.6617%	0.076162	接受

$$\alpha=0.05, K-S \text{ 臨界值}=0.2137$$

由上表的指標 I 可知，本研究模式相較古典引力模式皆減少變異。

表 5-2 模式間小抽樣比例樣本推估程度比較

	ERR	I	K-S 統計量	決策
古典引力模型	28490	-----	0.125529	接受
模式(一)	26923	-5.5001%	0.091169	接受

$\alpha=0.05, K-S$ 臨界值=0.2137

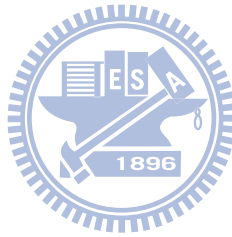
由上表的指標 I 可知，當以小抽樣比例的樣本推估時(每 30 比抽一比的樣本)，本研究模式相較古典引力模式減少變異。

由上述結果可知，本研究模式比古典引力模式推估程度較好。並以下列指標作為評估模式差異：

$$I = \frac{ERR_2 - ERR_1}{ERR_1} \times 100\%$$

其中 ERR_1 表古典引力模式推估資料與母體的變異

ERR_2 表本研究模式(一)(二)推估資料與母體的變異



第六章 不完整資料插補

6.1 引力模式插補不完整資料

不完整資料之種類可分為結構性與抽樣造成的缺失；由資料整理過程中，可發現當母體本身資料數量小，或是抽樣比例小時，會造成抽樣上的不完整資料。

本研究中，因為樣本資料即是實際旅次的部分百分比，已反應出母體的實際狀況，因此，以引力模式作不完整資料的微調插補。嘗試以模式(一)作插補。

模式(一):

引力模式

$$T_{ij} = W_1'(i)W_2(j)A_c(i)B_c^r(j)\exp[-\theta_c c_{ij}]$$

其中

$$W_1'(i) = \left[\sum_{g \in I} W_2(j)B_c^r(j)\exp[-\theta_c c_{ij}] \right]^{-1}$$

$$W_2(j) = B_c^r(j) \left[B_c^r(j) \sum_i W_1'(i)A_c(i)\exp[-\theta_c c_{ij}] \right]^{-1}$$

模式(一)為指數型引力模式。

6.2 不完整資料插補流程

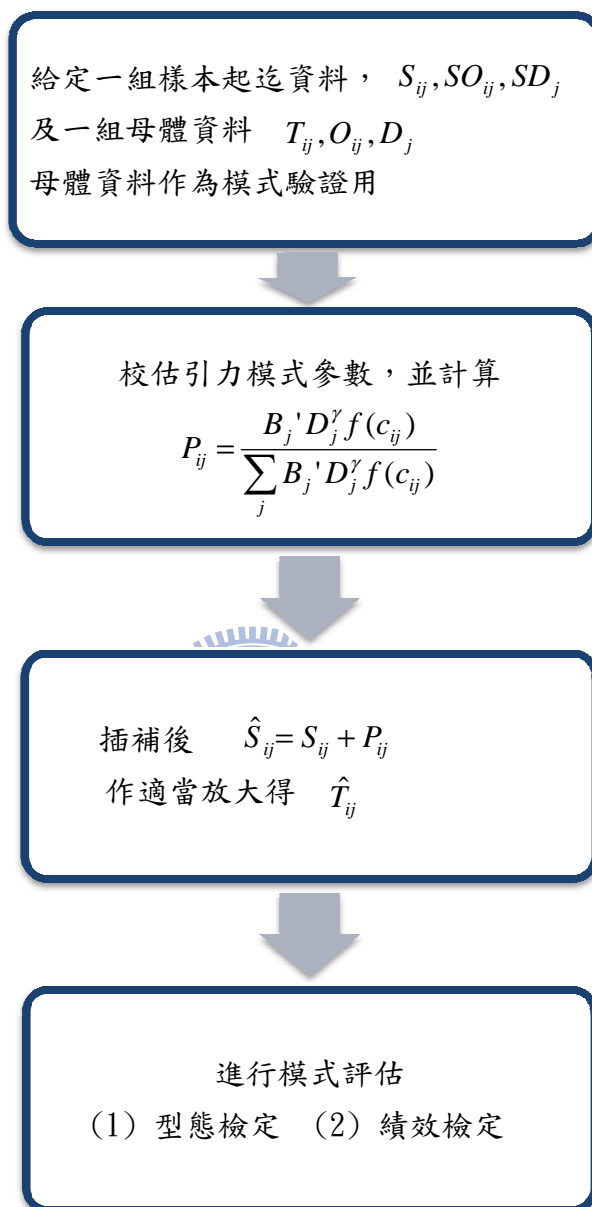


圖 6-1 不完整資料插補流程圖

[資料來源:本研究整理]

6.3 不完整資料處理

在本節中擬以修正 Deming Stephan iteration proportional fitting procedure 模式(Jou,Y. J., et al.,2003；以 DS 簡稱)與本研究模式(以 GM 簡稱)作比較，結果如下表所示：

	ERR	I	樣本數
GM(2) 註1	24293.21	-0.1540%	20847
DS(2)	24330.67	-----	20847
GM(10) 註2	25747.04	-0.3760%	1588
DS(10)	25844.21	-----	1588
GM(20) 註3	29610.80	-38.2640%	470
DS(20)	47963.58	-----	470
GM(30) 註4	30300.76	-40.5299%	235
DS(30)	50951.27	-----	235

母體總數:44050

註 1: (2)表每 2 筆抽一筆的樣本

註 2: (10)表每 10 筆抽一筆的樣本

註 3: (20)表每 20 筆抽一筆的樣本

註 4: (30)表每 30 筆抽一筆的樣本

$$\text{其中 } I = \frac{ERR_{GM} - ERR_{DS}}{ERR_{DS}} \times 100\%$$

由第五章可得引力模式在旅次分配推估上已有良好的表現，可知引力模式對旅次資料有良好的解釋力，根據上述結果亦可得到引力模式可針對本類型資料做不完整資料插補處理。

第七章 結論與建議

7.1 結論

本研究主要動機在於:希望提供旅次起迄分佈調查藉由可獲得且有限的資料,推估有效的高速公路車輛起迄分佈調查資料。研究所得的資訊歸納如下:

1. 根據理論基礎,將抽象的起迄點之間距離關係,利用校正因子、參數來平衡、量化為具體的數學式,本研究提出引力模式可作為高速公路交通旅次推估。
2. 在資源有限的情況下,若要進行旅次起迄分佈調查,可藉由簡易可獲得的交通 O_i, D_j 的資訊進行,本研究證明提出的模式確實能有良好的旅次推估。
3. 在人力、物力及經費有限的情況下,利用抽樣比例大的樣本進行旅次推估,本研究證明仍能有良好的推估能力。
4. 由於引力模式在旅次分配推估上已有良好的表現,根據實驗結果亦可得可作為本類型資料不完整資料插補處理。

7.2 建議

高速公路車輛起迄分佈調查資料，是運輸規劃中基本且重要的一環，對於這方面資料的取得，除了傳統調查方法外，國內外學者亦嘗試以可獲得的資料代入模式處理的方式來推估。在校估參數過程，除了利用窮舉法或是迭代方式獲得之外，可採取數學方法，如：牛頓法、梯度尋求法，快速獲得解，且可縮短校估所需時間。

由於抽樣而得的旅次起迄資料已隱含各區間活動強度與社經資訊，未來可考慮除了起、迄量與路段資訊外，尚可加入其他限制式，使模式更接近實際狀況。

引力模式校估出的參數，代表著各變量對於 T_{ij} 的敏感度，可嘗試分析各組參數與 T_{ij} 的關係，期望能藉由參數分析解釋運輸需求中的潛在需求。再者，本研究所驗證的資料係屬於高速公路主線資料，若未來可獲得主線及匝道資料，可將資料代入模式中進行推估，並藉由分析參數相關來作高速公路上交通量分析。

參考文獻

- [1] Ashish, S. and T.E. Smith., *Gravity models of Spatial Interaction Behavior*, Springer-Verlag, Tiergartenstrasse 17, D-69121, Heidelberg, Germany
- [2] Alemeida, M. W. and Goncalves, M. B., “A methodology to incorporate behavioural aspects in trip-distribution models with an application to estimate student flow, *Environment and planning A*, vol. 33, 2001.
- [3] Bell, M.G.H., “The estimation of origin–destination matrices by constrained generalized least squares.” *Transportation Research, Part B*, vol. 23, pp. 257–265, 1991.
- [4] Cascetta, E., “Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator.” *Transportation Research, Part B*, vol. 18, pp. 289–299, 1984.
- [5] Goncalves, M. B., ”The development of a new gravity-opportunity model for trip distribution”, *Environment and planning A*, vol. 25, pp.517-826,1993.
- [6] Goncalves, M. B., ”A comparative analysis of some trip distribution models used to estimate passenger flow”, *Anais do VII congress Nacional de Ensino e Pesquisa em Transporte*, vol. 1, pp.337-8348,1993.
- [7] Huff, D.L., “A probability analysis of shopping center trade areas”, *Land economics*, vol. 39, pp. 81-90, 1963.
- [8] Huff, D.L., “Definite and estimating a trading area”, *Journal of marketing*, vol. 28, pp. 34-38,1964.
- [9] Huff, D.L., *Parameter Estimation in the Huff Model*, ArcUser, 2003.
- [10] Lakshmanan, J. R. and W. G. Hansen.,” A retail market potential model”, *Journal of the American Institute of Planners*, vol.31, pp.134-43,1965.
- [11] Jou, Y.J., et al., “The Study of Origin-Destination Sampling Survey in Highway and Incomplete Data Analysis.” *Transportation Planning Journal*, Vol. 25, No. 4, pp709-726, 1996.
- [12] Jou, Y. J., et al., “Incomplete Information Analysis for the Origin-Destination

- Survey Table”, *J. Urban Plng. and Devel.*, Vol.132, Issue 4, pp. 193-200, 2006.
- [13] Ortuzar, J. de D. and L. G. Willumsen., “Modeling transport”, *John Wiley & Sons*, New York, 2001.
- [14] Reilly, W. J., “The law of retail gravitation”, *G. P. Putman and Sons*, New York, 1931.
- [15] Roy, J. R., “Comment on the development of a new gravity-opportunity model for trip distribution”, *Environment and planning A*, vol.25, pp.1689-91, 1993.
- [16] Stouffer, S. A., “Intervening opportunities: a theory relating mobility and distance”, *American sociological Review* 5, vol.5, pp.845-867, 1940.
- [17] Stouffer, S. A., “Intervening opportunities and competing migrants”, *Journal of Regional science association*, vol.5, pp.47-56, 1960.
- [18] Wootton, H. J., Calibration a gravity model and estimating trip ends from a partially observed trip matrix. Memorandum of December 21, 1972.
- [19] Wilson, A.G., *Urban and regional models in geography & planning*, Pion, London, 1967.
- [20] Wilson, A.G., ”A statistical theory of spatial distribution models”, *Transportation Research*, Vol. 1, pp. 253-269
- [21] Wilson, A.G., *Entropy in urban and regional modeling*, Pion, London, 1970.
- [22] Wills, M. J., ”A flexible gravity-opportunity model for trip distribution”, *Transportation Research B*, vol.20, pp.89-111, 1986.
- [23] 卓訓榮、周幼珍等，「公路車輛起迄調查方法之研究」，期末報告，交通部運輸研究所委託，民 84。
- [24] 李霞，「公路車輛起迄抽樣調查與不完整資料處理之研究」，國立交通大學交通運輸研究所，碩士論文，民 85。
- [25] 卓訓榮，「都市空間細分分派模型之研究—以台北市為例」，國立中興大學都市計劃研究所，碩士論文，民 69。

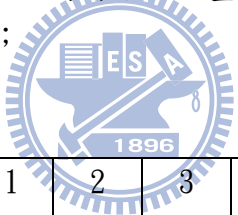
附錄、一

利用本研究所整理的起迄點流量資訊，計算線性路網路段流量，以4×4的旅次矩陣為例：

	1	2	3	4
1	0	T_{12}	T_{13}	T_{14}
2	T_{21}	0	T_{22}	T_{24}
3	T_{31}	T_{32}	0	T_{42}
4	T_{41}	T_{42}	T_{43}	0

圖附-1 站 1 至站 2 間的路段流量

在線性路網中，介於站 1 至站 2 間的路段流量為 $T_{12}, T_{13}, T_{14}, T_{21}, T_{31}, T_{41}$ 交通量總和，即為上圖附-1 灰色交通量；



j	1	2	3	4
1	0	T_{12}	T_{13}	T_{14}
2	T_{21}	0	T_{22}	T_{24}
3	T_{31}	T_{32}	0	T_{42}
4	T_{41}	T_{42}	T_{43}	0

圖附-2 站 2 至站 3 間的路段流量

介於站 2 至站 3 間的路段流量為上圖附-2 灰色交通量之總和；

圖附-3 站 3 至站 4 間的路段流量

j	1	2	3	4
1	0	T_{12}	T_{13}	T_{14}
2	T_{21}	0	T_{22}	T_{24}
3	T_{31}	T_{32}	0	T_{42}
4	T_{41}	T_{42}	T_{43}	0

介於站 3 至站 4 間的路段流量為上圖附-3 灰色交通量之總和；

[資料來源:本研究整理]

因此，關於路段資訊：

可知，當把所有路段流量總和即為：

$$\sum_k f_k = 2 \sum_i \sum_j T_{ij} - s$$

其中 $s = \sum_m T_{m,m+1} + \sum_n T_{n+1,n}$

f : 路段流量

上述的路段資訊，為推導模式(二)用。