

規律折扣數列與高齡死亡率

王信忠 *

余清祥 **

摘要

自 20 世紀中葉以來，人類平均壽命屢創歷史紀錄，高齡（65 歲以上）人口在許多國家已經或即將超過全國人口的五分之一，老年人成爲 21 世紀的熱門研究議題，但由於高齡人口資料在 1990 年代後才有較完整紀錄，至今對於何種高齡死亡率模型較佳仍無定論。本文引入吃角子老虎問題(Bandit Problem)的規律折扣數列(Regular Discount Sequence)，用來描述老年人的平均餘命變化，以及預測未來的高齡死亡率。許多常用的死亡率模型，例如：Gompertz law、均勻死亡(Uniform Distribution of Death)假設、Constant Force、以及 Hyperbolic 假設等都滿足規律折扣數列的條件。

除了理論推導，我們採用美國加州大學柏克萊分校(University of California, Berkeley)的 HMD (Human Mortality Database)資料，包括日本、美國及臺灣的死亡率資料，驗證規律折扣數列，三個國家的生存數與平均餘命均大致符合折扣數列的假設。另外，我們也使用布朗運動(Brownian Motion)隨機微分方程式，建立折扣數列的模型，用來預測未來的高齡人口死亡率，電腦模擬顯示無論是數列比值或是死亡率的預測，折扣數列模型都有不錯的結果，亦即本文提出的模型可用於預測高齡死亡率。

關鍵字: 吃角子老虎問題、布朗運動、長壽風險、蒙地卡羅模擬、區塊拔靴法

*真理大學統計與精算系助理教授， au4369@mail.au.edu.tw，台北縣淡水鎮真理街 32 號， 02-2621-2121 ext 3311

**政治大學統計系教授， csyue@nccu.edu.tw，台北市文山區指南路二段 64 號， 02-29387695

Using Regular Discount Sequence to Model Elderly Mortality

Hsin Chung Wang

Jack C. Yue

Abstract

Life expectancies of the human male and female have been increasing significantly since the turn of 20th century, and the trend is expected to continue. The study of the elderly mortality is thus become one of the most favorite research topics. However, because there were not enough elderly data before 1990, there is still no conclusion about which mortality model is appropriate for describing the elderly mortality. In this study, we modify the regular discount sequence in Bandit Problem and use it to describe the elderly mortality. We found that many frequently used mortality models, such as Gompertz Law, and famous mortality assumption (Uniform Distribution of Death, Constant Force, and Hyperbolic assumption) all satisfy the requirement of regular discount sequence.

We also use empirical data from the HMD (Human Mortality Database from University of California, Berkeley), including data from Japan, US, and Taiwan, to evaluate the proposed approach. The discount sequences of life expectancy and surviving number ratio do satisfy the regularity condition. In addition, we use the Brownian Motion Stochastic Differential Equation to model the discount sequence. Using this model, we predict the future mortality rates and life expectancy. The simulation study shows some promising results.

Key words: Bandit Problem, Brownian Motion, Longevity Risk, Monte Carlo Simulation, Block Bootstrap

壹、前言

除了經濟及社會發展的進步外，近年臺灣地區居民壽命也趕上先進國家，堪稱已開發國家的一員¹，若以 65 歲人口比例超過 7% 作為高齡化社會的標準，臺灣在 1993 年首度突破後，2009 年底更超過 10.5%。圖 1 及圖 2 中為臺灣、日本與美國在 1970~2005 年的比較，各國在男、女兩性零歲與 65 歲時平均餘命均呈現線性成長趨勢，其中又以臺灣女性成長幅度最大（資料來源：HMD，Human Mortality Database from University of California, Berkeley）。依據行政院經濟建設委員會（以下簡稱：經建會）在 2010 年公佈的推估報告（資料來源：經建會「2010 年至 2060 年臺灣人口推計」），民國 114 年時 65 歲以上老年人口將超過全國人口 20%，人口老化速度遠高於歐美國家，未來臺灣地區人口老化情形將更趨明顯。

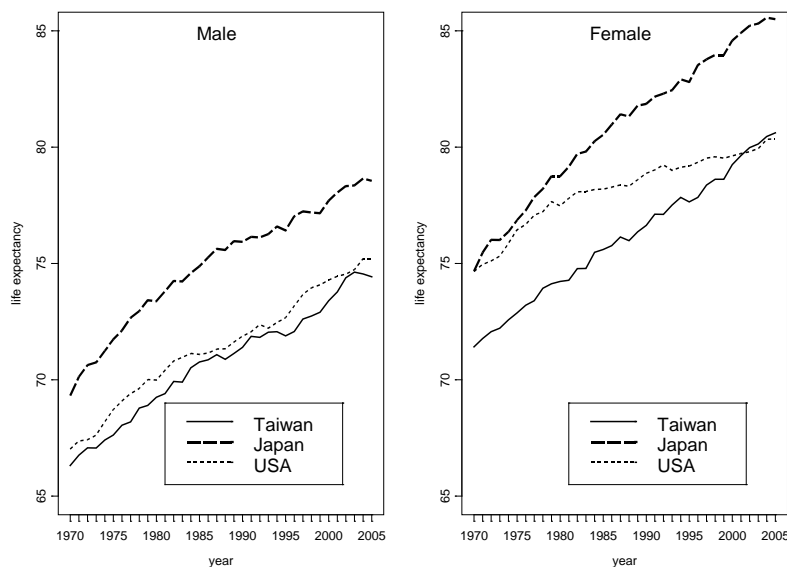


圖 1、各國 1970~2005 年零歲平均餘命

另由內政部統計處資料，臺灣 2010 年的總生育率為 0.9，也就是平均每名婦女生育不到 1.0 個小孩，低於日本的 1.4²，少子化情形相當嚴重。從人口老化與少子化的現象顯示，傳統家庭子女扶持老人功能將漸趨式微，未來老年人的經濟安全，將有賴於社會保險制度與私人保險規劃。而無論是政府

¹根據 2011 年內政部統計處資料，臺灣地區居民的平均壽命約比美國多一歲，比德國約少一歲。

²行政院經濟建設委員會，民國 98 年 8 月 28 日新聞稿，2009 年世界人口重要指標。

社會保險及民間商業保險，費率計算都仰賴較為可信的死亡率模型，確保制度能夠永續經營、以及保險公司的財務自主，近年各國學者依據各國特性提出不同模型，以符合各地死亡率的變化趨勢。

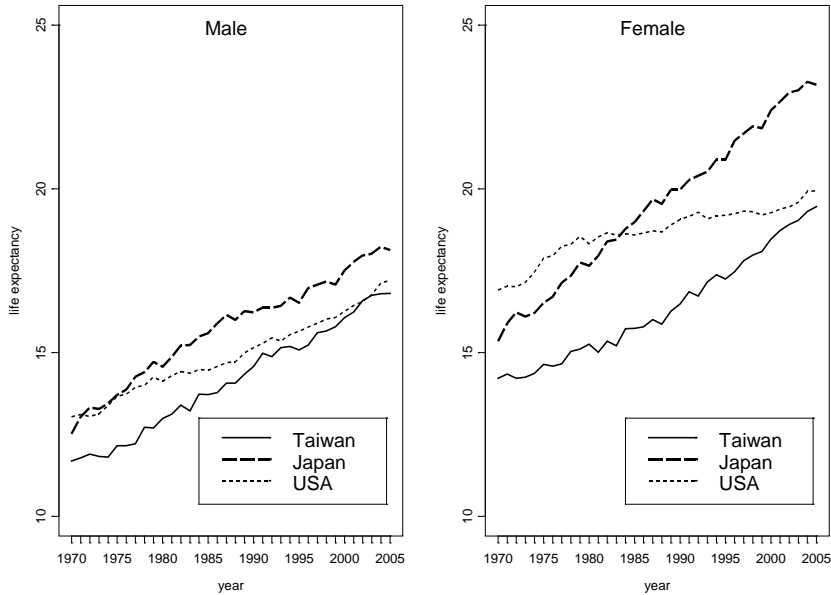


圖 2、各國 1970~2005 年 65 歲平均餘命

有關死亡率模型的研究，最早可追溯至 1662 年，當時 Graunt 探討 1603~1661 年倫敦的出生率與死亡率，並以抽樣的方法來估計人口；1693 年 Halley 根據 Breslau 居民死亡資料，發展出全球第一個生命表，建立死亡率與年齡間的關係，發現兩者存有高度相關。人口學家通常認為死亡率與時間有關，將年齡、時間視為死亡率模型中的參數，用函數模型解釋死亡率的變化。後續如 1729 年 De Movire 提出生存曲線模型、1825 年 Gompertz 模型、1860 年 Makeham 模型、1939 年 Weibull 模型、1980 年 Heligman and Pollard 模型等，都屬於靜態確定型模型，藉由歷史資料估計模型中的參數以描述死亡率，但沒有將隨著時間參數演變的死亡率改善列入考量。

因為死亡率預測的需求日殷，加入時間參數變動的死亡率模型應運而生，例如：Lee and Carter (1992) 的隨機死亡率模型，因為操作方便、模型配適能力不錯，目前廣為各國學者及政府採用。這個模型假設中央死亡率滿足

$$\ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t}$$

上式中的 $m_{x,t}$ 代表 x 歲在 t 年的中央死亡率，以及三個參數： α_x 為 x 歲死亡率的平均曲線， β_x 為 x 歲相對死亡率的變化率， κ_t 是 t 年死亡率強度的變化量，誤差 $\varepsilon_{x,t}$ 為常態分配。Lee and Carter 模型的特色之一是死亡率的對數隨時間下降，許多國家的死亡率滿足 κ_t 是時間 t 的線性函數。

1992 年以來，不少研究針對 Lee-Carter 模型提出修正模型。Chen and Cox (2009) 考量重大災難等引起死亡率跳動，以幾何布朗運動(Geometric Brownian Motion)與複合卜瓦松過程(Compound Poisson Process)估計 κ_t 。Renshaw and Haberman (2006) 在原有的時間因子 κ_t 外，再加入一個世代因子；Yang et al. (2010) 則以主成分分析(Principal Component Analysis; PCA)，加入第二個時間因子。Cairns et al. (2006a、2006b) 以殖利率模型考量死亡率，提出短期死亡率模型、遠期死亡率模型、市場死亡率模型與兩因素廣義模型等。

或許各國國情不同，可用的高齡死亡率資料較少，至今尚無死亡率模型得到壓倒性的支持，而臺灣地區則受限於 85 歲以上資料的不足，產官學各界對於死亡率的處理方式也是大不相同。像是內政部統計處編算生命表時，高齡死亡率推算目前採用高馬氏模型(Gompertz-Makeham)加權迴歸修勻法；保險業則偏好 Lee-Carter 模型，參數估計使用奇異值分解(SVD; Singular Value Decomposition)。

本文引進吃角子老虎問題(Bandit Problem)的折扣數列，驗證常見的生命表函數(例如：死亡率、平均餘命)是否服從折扣數列規律性，測試這個模型用於死亡率推估，作為另一種可能的高齡死亡率模型。本文編排如下：第二節介紹折扣數列的定義及想法，並說明其與死亡率的關聯，並驗證幾個常見的死亡率模型滿足折扣數列的規律性；第三節以 HMD 中之日本、美國與臺灣的經驗資料驗證折扣數列的規律性，包括生命表中的各年齡平均餘命、死亡數、生存數與死亡率等數值的歷年趨勢；第四節則以布朗運動方法估計折扣數列，透過蒙地卡羅模擬驗證估計結果，再與 Lee-Carter 模型比較臺灣死亡率的預測表現；第五節則為結論與建議，提出本文研究的一些想法與後續研究的可能方向。

貳、折扣數列與死亡率

所謂折扣數列(Discount Sequence)，源自於吃角子老虎問題，就是在賭場玩吃角子老虎（又名獨臂大盜）若干次，如何選擇報酬最佳的吃角子老虎。折扣數列就是用來規範玩的次數，折扣數列的規律性指的是數列具有漸增失敗率(Increasing Failure Rate)，在下一段詳細定義及討論。如果折扣數列滿足規律性，尋找最佳化策略會比較容易，例如：Berry and Fristedt (1985)及 Yue (1999)證明雙臂吃角子問題中，折扣數列若具有規律性，則選擇報酬變異數較小的那一臂會是最佳決策。

本文將以折扣數列與存活函數連結，說明死亡率的性質，首先介紹吃角子老虎問題與折扣數列。

常見的假設為 k ($k \geq 2$) 臂吃角子老虎問題，要從有 k 個手臂(arms)的吃角子機器(slot machine)，選擇一個手臂進行拉霸(play)來獲得報酬，每次只能拉一個手臂，每次拉霸可獲得隨機而定的報酬。因為變異數與樣本數成反比，嘗試同一手臂愈多次、資訊愈完整，但若嘗試太多次投資報酬較差的手臂，卻不利於整體報酬。換言之，我們必須在累積資訊與獲得獎勵(Information vs. Immediate Payoff)中取得平衡。吃角子老虎問題早期應用於臨床實驗(Clinical Trial)上，假設某一疾病有 k 種治療方式，目標是治癒愈多病患，而非犧牲某些病患找到最佳處方，必須在嘗試藥效較不清楚的新處方，或是繼續使用現有處方間取得平衡。

假設我們有二台吃角子老虎 R 及 L ，稱為右臂及左臂，每拉霸一次、產生一個觀察值，目的是最大化 $\sum_{m=1}^{\infty} \alpha_m Z_m$ 的期望值，其中 Z_m 值第 m 次的觀察值。定義 $A = (\alpha_1, \alpha_2, \dots)$ 為折扣數列，它是非遞增的非負數列 ($\alpha_1 \geq \alpha_2 \geq \dots \geq 0$)，滿足 $\sum_{m=1}^{\infty} \alpha_m < \infty$ ，而隨機變數 $\sum_{m=1}^{\infty} \alpha_m Z_m$ 即為拉霸的報酬。

在吃角子老虎問題中，根據 Berry and Fristedt 的定義：如果對每一個觀察值 n ， $\gamma_n = \sum_{i=n}^{\infty} \alpha_i$ 滿足

$$\gamma_n \cdot \gamma_{n+2} \leq (\gamma_{n+1})^2 \quad (1)$$

則此折扣數列 $(\alpha_1, \alpha_2, \dots)$ 為規律(Regular)數列。

將折扣數列套用至生命表函數，亦即以生存數等函數代入 γ_n 檢驗(1)式，其中生存數與平均餘命隨著年齡遞減，因此滿足折扣數列的規律性；死亡率與死亡數在高齡組大致滿足規律性，但在幼齡、青少年及成年等階段不盡然都是遞降數列，未必符合規律性。本文的研究目標為高齡人口死亡率，並以折扣數列找出生命表函數的遞降幅度，捕捉這些數值在各年齡間的變動情形，除非另外說明，以下的討論將僅考慮死亡率及死亡數。

折扣數列的規律性也具有數學及統計上的意義。將折扣數列比值取對數後，符合規律性的折扣數列本身也滿足對數凹函數(log concave function)，因為二次微分小於0，通常可藉由微分找到函數的最大值。在機率理論裡，如果機率密度函數(density function)是對數凹函數，它的累積分配函數(cumulative distribution function)及存活函數(survival function)也是對數凹函數。對數凹函數可應用於不同領域，例如在財務領域裡，可以透過二次微分找到獲取最大利潤的投資組合；在保險精算領域裡，找出提存最多準備金的時間點保險公司可避免現金不足的問題。本文的用意在於透過對數凹函數的特性，提供高齡人口死亡率的參考依據，用以推估未來高齡人口的死亡率。

為了建立規律折扣數列與死亡率的關聯，將折扣數列隨機變數 N 視為存活時間 (Survival Time) T ，定義 $\gamma_n = P(N \geq n)$ ，亦即令 γ_x 為存活函數 $S(x) = P(T \geq x)$ ，可以推導出(1)式與下列(2)式相同：

$$l_n \cdot l_{n+2} \leq (l_{n+1})^2 \quad \text{或} \quad \frac{l_n \cdot l_{n+2}}{(l_{n+1})^2} \leq 1, \quad (2)$$

其中 l_n 是生命表中年齡 n 歲的生存數。若給定生命表的基數(Radix) l_0 ，通常定義為 100,000 人，則方程式(2)也可記為 $\frac{p_{n+1}}{p_n} \leq 1$ 或 $q_{n+1} \geq q_n$ ，即若存活時間滿足規律性，相當於死亡率是年齡的非遞減函數。上述 p_x 為現年 x 歲居民存活至 $x+1$ 歲的機率，記為 $p_x = P(T > x+1 | T > x)$ ，因此生存數 l_x 滿足 $l_x = l_0 \times p_x$ ； q_x 為現年 x 歲居民在 $x+1$ 歲前的死亡機率，可表為 $q_x = P(x < T < x+1 | T > x) = 1 - p_x$ 。

若以生命表的定常人口(Stationary Population)表示，上述的方程式也可表為下列(3)、(4)、(5)的格式，亦即驗算折扣數列是否滿足規律性時，可以轉換成以生存數或平均餘命函數的方程式來計算，至於其他函數方程式以模擬實證的結果，發現在高齡的部分也能滿足。下列三式中的 ${}^o e_x$ 為x歲的平均餘命，或是現年x歲居民未來可再存活的時間； d_x 為x歲至x+1歲的死亡數，可以透過生存數及死亡機率表示，記為 $d_x = l_x \times q_x = l_x - l_{x+1}$ 。

$${}^o e_n \cdot {}^o e_{n+2} \leq ({}^o e_{n+1})^2 \quad \text{或} \quad \frac{{}^o e_n \cdot {}^o e_{n+2}}{({}^o e_{n+1})^2} \leq 1. \quad (3)$$

$$d_n \cdot d_{n+2} \leq (d_{n+1})^2 \quad \text{或} \quad \frac{d_n \cdot d_{n+2}}{(d_{n+1})^2} \leq 1. \quad (4)$$

$$q_n \cdot q_{n+2} \leq (q_{n+1})^2 \quad \text{或} \quad \frac{q_n \cdot q_{n+2}}{(q_{n+1})^2} \leq 1. \quad (5)$$

除了以生命表中的生存數、平均餘命、死亡數、死亡率，探討吃角子老虎問題的折扣數列規律性，如果滿足定常人口假設，建議也可直接藉由檢查方程式(5)的死亡率 q_n ，而不必先編修生命表。但因為高齡人口數較少，死亡率的震盪幅度較大，預期方程式(5)的死亡率比值也會有較大的震盪。

除了定義外，以下探討常用的幾個死亡率模型（包括非整數年齡的死亡假設），檢查是否滿足折扣數列的規律性。為簡化表示，以下討論定義折扣數列比值 $r_n = \frac{\gamma_n \cdot \gamma_{n+2}}{(\gamma_{n+1})^2}$ ，只要比值不大於一即滿足規律性， γ_n 原則上以生存數（或存活函數）及平均餘命為主。

一、Gompertz 模型

人口老化是近年世界各國的共同趨勢，但由於缺乏較完整的高齡死亡率資料，Gompertz 法則仍然是廣為使用的高齡死亡模型之一，這個模型假設瞬間死亡率隨隨年齡而指數上升：

$$\mu_x = BC^x, \quad B > 0, C > 1, \quad (6)$$

其中 x 表年齡， μ_x 為瞬間死亡率(Instantaneous Mortality Rate)或死力(Force of Mortality)，由存活函數微分而得，或是 $\mu_x = \frac{S'(x)}{S(x)}$ 。因為存活機率滿足

$${}_t p_x \equiv P(T > x+t | T > x) = \exp\left(-\int_0^t \mu_{x+s} ds\right) = \exp\left(-\frac{BC^x}{\log(C)}(C^t - 1)\right),$$

因此 Gompertz 法則與下式(7)等價

$$\frac{\log(p_{x+1})}{\log(p_x)} = C. \quad (7)$$

如果以中央死亡率 m_x 代替 μ_x ，則 $k_{x+1} \equiv \log(m_{x+1}/m_x) = \log(C)$ 為一常數。給定

$$\mu_x = BC^x, \quad \text{並且因為 } C > 1, \quad \text{可以得到 } \gamma_n = {}_n p_0 = \exp\left(-\frac{B}{\log(C)}(C^n - 1)\right) \text{ 及}$$

$$\frac{\gamma_n \cdot \gamma_{n+2}}{(\gamma_{n+1})^2} = \exp\left(-\frac{BC^n}{\log(C)}(C^2 - 2C + 1)\right) = \exp\left(-\frac{BC^n}{\log(C)}(C - 1)^2\right) < 1.$$

因此如果死亡率是服從 Gompertz 法則，則(1)式成立，即折扣數列為規律數列。

同樣地，如果死亡率是服從 Makeham law 即 $\mu_x = A + BC^x$ ，(1)式的規律性質也一樣滿足。

另外，生命表中一般只標示整數年齡的數值，非整數年齡的死亡率多以三種假設求得：均勻假設、定死力假設、雙曲線假設，我們將逐一這三種非整數年齡假設是否也滿足規律性。

二、均勻死亡(Uniform distribution of death; UDD)假設

死亡率在 UDD 假設下，則對於任意的 $t, m, 0 \leq t \leq m$ ，滿足

$$l_{n+t} = \frac{m-t}{m} \cdot l_n + \frac{t}{m} \cdot l_{n+m}.$$

則(2)式等價於 $\frac{l_{n+2}/l_{n+1}}{l_{n+1}/l_n} \leq 1$ ，或 $p_{n+1} \leq p_n$ 。即 UDD 假設滿足(2)式規律性。

三、定死力(Constant Force; CF)假設

假設死力是常數，即對所有年齡 x ， $\mu_x = \mu$ 。那麼 ${}_n p_x = e^{-n\mu} = \frac{l_{x+n}}{l_x}$ 並且

$$\frac{l_n \cdot l_{n+2}}{(l_{n+1})^2} = \frac{e^{-(2n+2)\mu}}{e^{-(2n+2)\mu}} = 1。即定死力假設也滿足(2)式的規律性。$$

四、雙曲線(Hyperbolic)假設

若死亡率滿足雙曲線假設，則對於任意的 t 、 m ($0 \leq t \leq m$)，滿足

$$\frac{m}{l_{n+t}} = \frac{m-t}{l_n} + \frac{t}{l_{n+m}}, \text{ 或 } l_n \cdot l_{n+2} = l_{n+1} \cdot \frac{l_n + l_{n+2}}{2}。如同 UDD 假設，可透過計算得$$

到 $l_{n+1} \geq \frac{l_n + l_{n+2}}{2}$ ，因此雙曲線假設也滿足 $l_n \cdot l_{n+2} = l_{n+1} \cdot \frac{l_n + l_{n+2}}{2} \leq (l_{n+1})^2$ 。即雙曲線假設滿足(2)式規律性。

本節我們驗證幾個死亡率假設及 Gompertz 等模型滿足規律性： $\frac{l_n \cdot l_{n+2}}{(l_{n+1})^2} \leq 1$

或 $p_{n+1} \leq p_n$ 。如果滿足此規律性，代表死亡率是年齡的非遞減函數，亦即死亡率隨著年齡遞增，死亡年齡隨著年代增加有愈往高齡集中的現象，理論上如果高齡死亡率函數，滿足對數凹函數特性，則其死亡率趨勢將不會是如指數函數般遞增。將對數凹函數應用於高齡人口死亡率的計算上，目前仍有些限制，因此我們將仿造 Coale and Kisker (1990) 的高齡死亡率編算，先以實證分析找出可能的趨勢，再結合電腦模擬方法建立模型。下一節將以經驗資料繼續驗證本文提出的方法。

參、實證資料分析與發現

我們在上一節介紹折扣數列及規律性，並且驗證幾個常見的死亡率模型滿足規律性，本節將繼續以實證資料驗證折扣數列的規律性，我們採用美國加州大學柏克萊分校的 Human Mortality Database (HMD) 資料庫。HMD 資料庫為美國加州大學柏克萊分校及德國 Max Planck 人口研究中心 (Max Planck Institute for Demographic Research) 兩個單位合作共同建立，蒐集全球 37 個國家的全國人口

資料，紀錄了每年的總人口數、死亡數、死亡率，以及各年齡的人口數、死亡數、死亡率，HMD是死亡率實證研究非常重要的資料庫³。除了檢驗臺灣地區的死亡率特性外，也考量日本、美國的資料，驗證折扣數列模型是否只適用於臺灣。實證分析時將資料分成三組：幼年年齡層(14 歲以下)、青壯年齡層(15-64 歲)與老年年齡層(65 歲以上)，觀察平均餘命、生存數之折扣數列比值變動的情形，並比較其歷年及各年齡間之差異。

另外，因為零歲及幼年人口的死亡率及死亡數有較大的變動，折扣數列比值當然也會有很大的震盪，並不適宜直接驗證規律性，本節將不討論；死亡率及死亡數折扣數列比值可參考附錄圖 1 及圖 2。不過，這三個國家的死亡率及死亡數在中、高年齡組都滿足規律性，未來可用於高齡人口死亡率，在下一節將介紹如何應用規律折扣於死亡率的預測。

一、平均餘命

為求資料年度的一致方便比較，我們使用 1970 年以後各國男、女性的生命表資料。以五齡組 0 歲、1-4 歲與 5-9 歲幼齡平均餘命折扣數列比值為例(圖 3-1)，比值都在 0.93 與 1.0 之間，其中除了臺灣的波動程度較大外，日本與美國比值均一致的隨著時間平緩增加，女性上升幅度較男性高。各國在幼年的平均餘命折扣數列，其比值均小於 1、且漸漸逼近於 1，都滿足規律性的要求。臺灣變化幅度較快且在幼年組有較大的震盪，猜測可能與人口數較少有關。

繼續探討青壯年的平均餘命折扣數列。以五齡組 30-34 歲、35-39 歲與 40-44 歲青壯年齡層為例(圖 3-2)，各國比值歷年震盪幅度均比幼齡組大，大約在 0.99 附近，大致也滿足規律性的要求。平均數命的比值以臺灣及日本的斜率較大，而且女性上升幅度也明顯比男性高，男性的比值卻有較大的震盪。

接著我們觀察高齡組 55-59 歲、60-64 歲、65-69 歲的平均餘命折扣數列比值(圖 3-3)，三個國家的比值都小於 1，滿足規律性的要求。美國近年的比值來都在 0.98 周圍變動，大於其他二國平均餘命折扣數列比值大約在 0.96~0.98 間變

³HMD 資料可免費下載，網址：<http://www.mortality.org/>。

化；以性別來看，則男性的比值普遍較女性高，表示男性年齡間平均餘命變動較大，這與青壯年及幼齡層明顯不同。分析發現，考量臺灣、日本、美國三國在1970年以來的生命表，無論是哪一個年齡組的平均餘命折扣數列比值，都滿足規律性的條件。

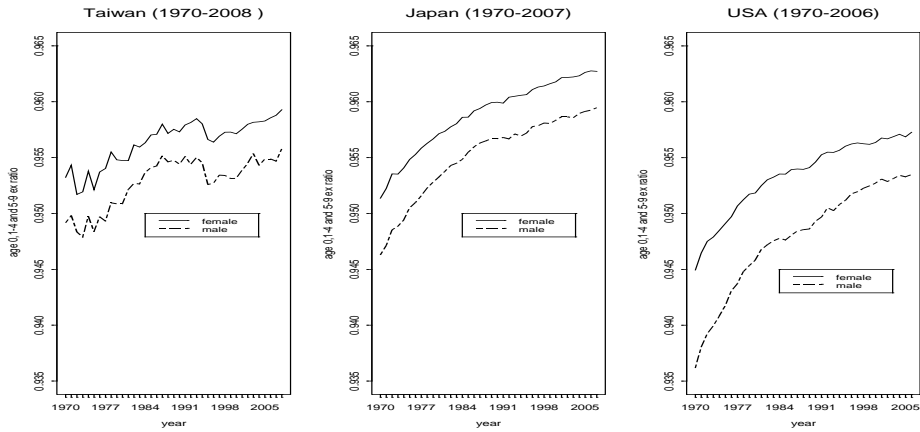


圖 3-1、幼年年齡層平均餘命折扣數列比值

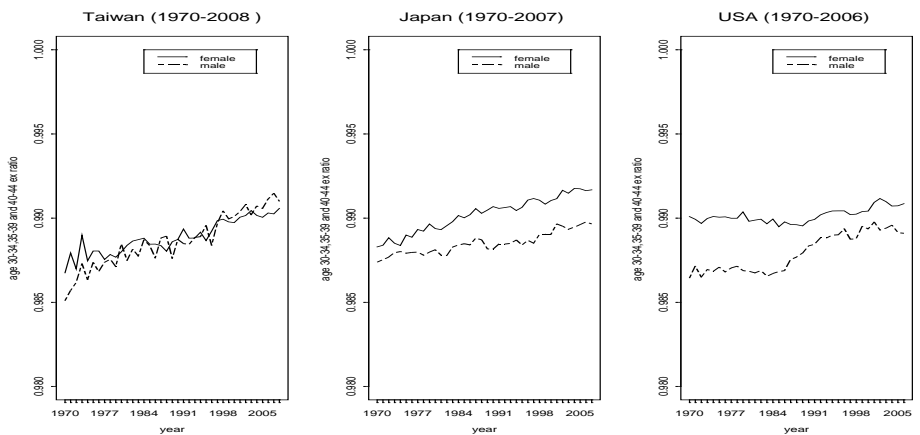


圖 3-2、青壯年年齡層平均餘命折扣數列比值

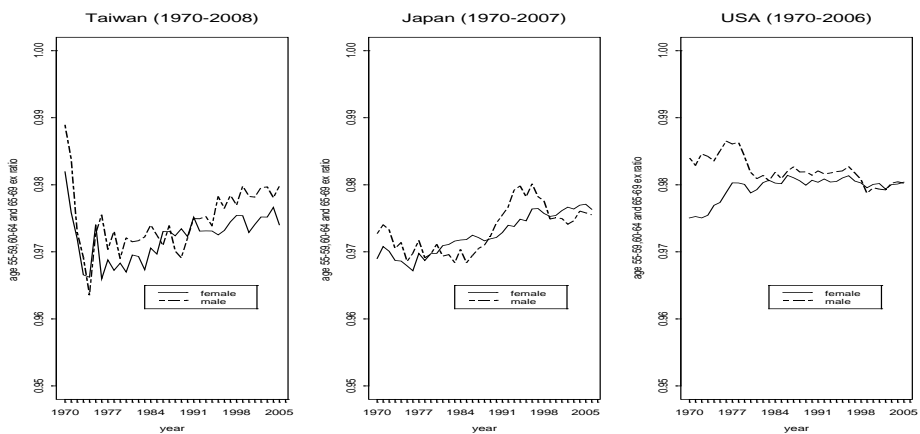


圖 3-3、老年年齡層平均餘命折扣數列比值

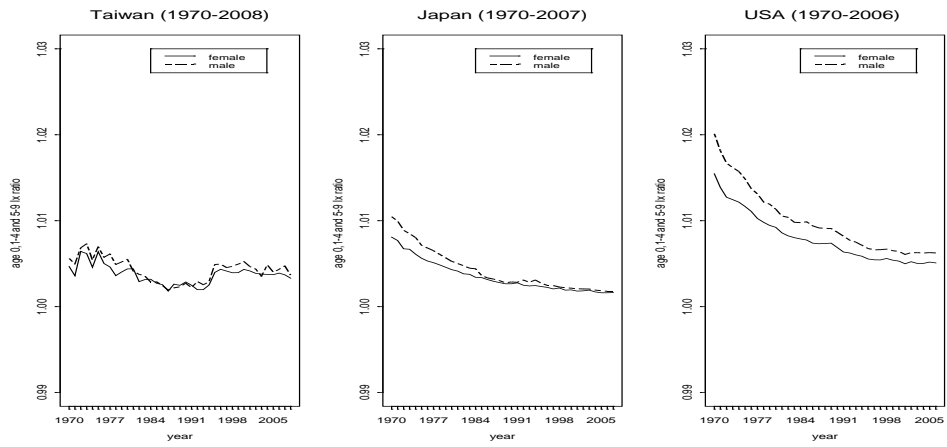


圖 4-1、幼年年齡層生存數折扣數列比值

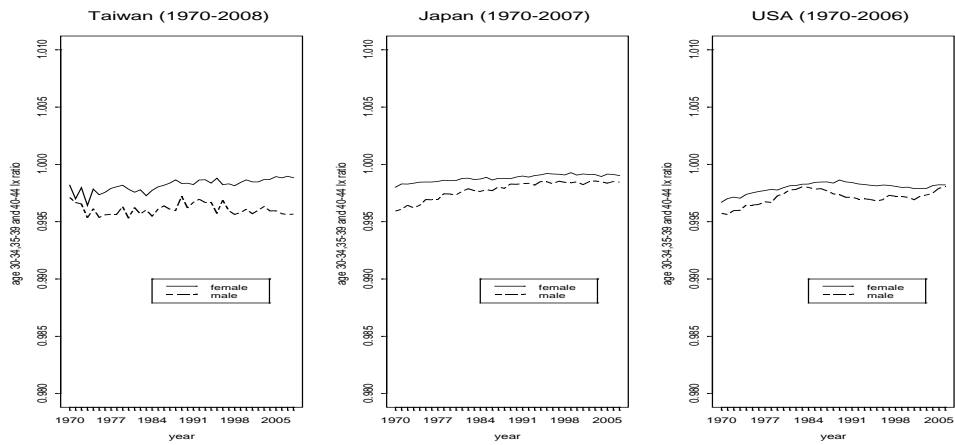


圖 4-2、青壯年年齡層生存數折扣數列比值

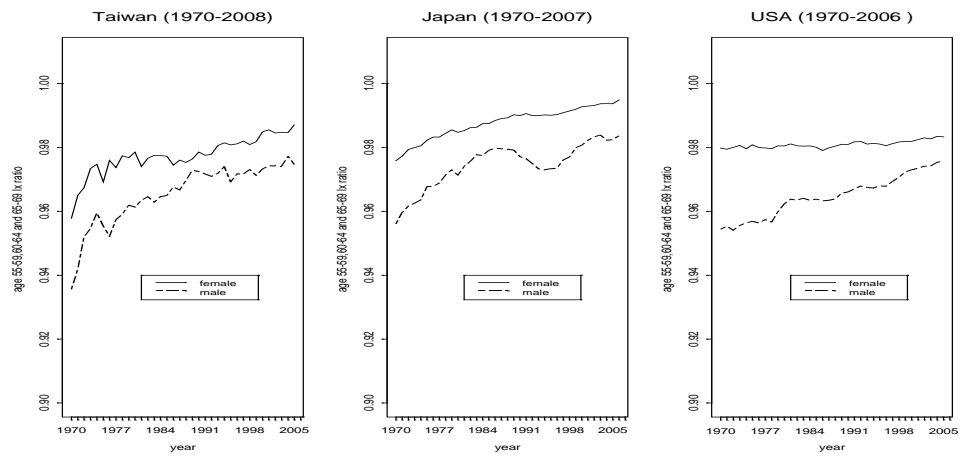


圖 4-3、老年年齡層生存數折扣數列比值

二、生存數

圖 4-1 為五齡組 0 歲、1-4 歲與 5-9 歲幼齡生存數折扣數列比值，數值雖然都大於 1，日本與美國的比值都漸漸收斂到 1，但臺灣的生存數比值則大約維持定值。這些結果顯示幼年人口的生存數比值，並不滿足(3)式的規律性要求。青壯年齡層 30-34 歲、35-39 歲與 40-44 歲的生存數比值在圖 4-2，比值的數值非常穩定，而且全部都小於 1；以性別來看，男性的比值普遍較女性高。高齡組的 55-59 歲、60-64 歲與 65-69 歲生存數比值參考圖 4-3，其數值都小於 1，而且很穩定地趨近 0.98；以性別來看，女性的比值均較男性高。綜合這些結果，我們發現青壯年及高齡組的生存數，均滿足折扣數列的規律性。

進一步觀察各國在不同年齡組的生存數盒型圖（圖 5），發現盒型圖有一個共同趨勢：比值隨著年齡呈現指數遞減，且在 20 歲以上的比值都小於 1。生存數的比值隨年齡遞減可用於推估死亡率改善幅度，進而推算各年齡的死亡率或平均餘命。例如：生存數折扣數列比值為 0.5，即 $r_x = \frac{l_{x+2}/l_{x+1}}{l_{x+1}/l_x} = 0.5 \leq 1$ ，表示 $\frac{l_{x+2}}{l_{x+1}} = 0.5 \frac{l_{x+1}}{l_x}$ 即 $q_{x+1} = 0.5 + 0.5q_x$ ，或是 $q_{x+1} = 1 - r_x + r_x q_x$ ，當我們知道 x 歲死亡率時， $x+1$ 歲以後的死亡率可經由疊代方法計算。另外，觀察各國最高年齡層生存數折扣數列比值的中位數，美國與臺灣大約是 0.6、日本約為 0.5，而且歷年美國與臺灣生存數折扣數列比值變動較大。

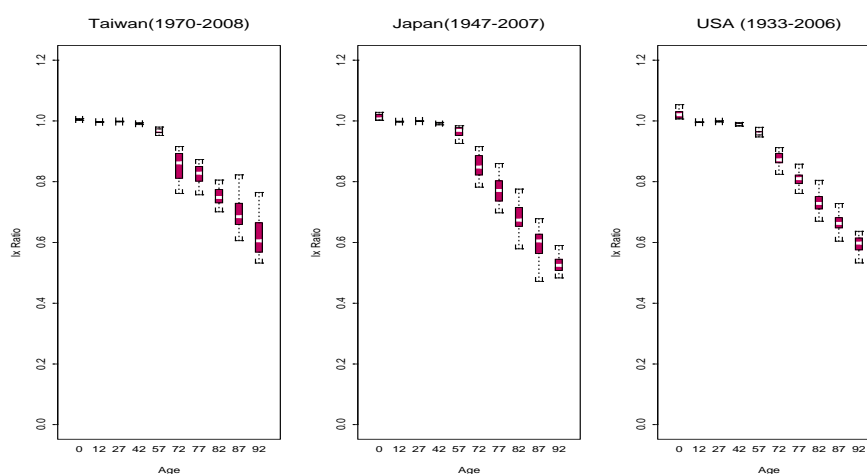


圖 5、各年齡組歷年生存數折扣數列比值盒型圖

以上檢驗生命表函數的折扣數列，平均餘命及生存數較能滿足規律性的要求，死亡率及死亡數則因震盪幅度較大，無法全部滿足規律性。因為生命表函數的折扣數列大致符合規律性，且比值大致穩定，可以作為推估未來死亡率之用，尤其是資料較缺乏、資料品質較難確認的高齡族群，可藉由青壯年族群的死亡率推算出高齡死亡率。下一節我們將提供以布朗運動(Brownian Motion)建立折扣數列比值的方法。

肆、模型配適與模擬應用

二十世紀以來，許多國家的各年齡死亡率對數值似乎隨時間線性下降，從 1960 年代的美國精算學會(SOA; Society of Actuaries)和現在廣為各國使用的 Lee-Carter 模型，都假設每個年齡層的對數死亡率可以由時間的線性函數表達，像是 Lee-Carter 模型通常可簡化為 $\ln(m_{x,t}) = \alpha_x + \beta_x \cdot t$ 的型態，其中 $m_{x,t}$ 是第 t 年 x 歲的中央死亡率， α_x 和 β_x 是年齡的函數。我們仿造這個想法，透過觀察折扣數列比值差分值隨著時間變化的情形，建立折扣數列比值的預測模型，以蒙地卡羅模擬方法來描述各年齡歷年死亡率的變動情形。

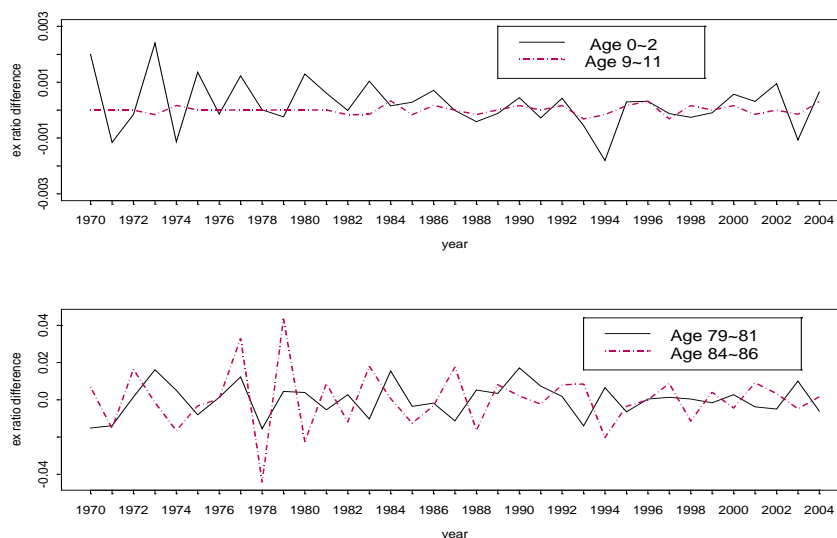


圖 6、臺灣男性 1970~2005 平均餘命 $\Delta r_x(t)$

一、折扣數列模型

考量死亡率具有不確定，若 $r_x(t)$ 為第 t 年 x 歲的平均餘命(或是生存數)折扣數列之比值，先計算 x 歲比值的差分值 $\Delta r_x(t) = r_x(t+1) - r_x(t)$ 。先以臺灣男性歷年的平均餘命為例，在此僅以四個年齡組為示範：0-2 歲、9-11 歲、79-81 歲、84-86 歲，觀察這四個年齡組比值差分值的變動趨勢(圖 6)。

由圖 6 可看到差分值具有隨機波動的現象，各年齡層的 $\Delta r_x(t)$ 在 0 附近上下震盪，以高年齡層的幅度較大，平均數大致接近 0。以 Wiener 過程的序列 $z(t)$ 表達比值的差分值，即 $\Delta r_x(t) = a\Delta t + b\Delta z(t)$ ，令 $\Delta t \rightarrow 0$ 可得 $dr_x(t) = adt + bdz(t)$ ，再結合常數趨勢項、震盪速度與波動現象，建立布朗運動死亡率隨機微分方程模型(Brownian Motion Mortality Stochastic Differential Equations Model：BMSDE)如下式：

$$dr_x(t) = k(\theta - r_x(t)) \cdot dt + \sigma \cdot dz, \quad dz \sim N(0, dt) \quad (8)$$

其中

k ：為 $r_x(t)$ 回到平均水準 θ 的力道強弱即為調整速度。

θ ： $r_x(t)$ 的長期平均水準。若 $r_x(t) > \theta$ ，則 $r_x(t)$ 大致會下降；若 $r_x(t) < \theta$ ，則 $r_x(t)$ 將上升。

σ ：擴散係數(Diffusion Coefficient)。

為使此一連續模型容易計算，以離散化(Discrete)的形式表達(8)式：

$$\Delta r_x(t) = k(\theta - r_x(t)) + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2)$$

可使用最小平方法估算參數 k 、 θ 、 σ ，參數估計值以 \hat{k} 、 $\hat{\theta}$ 、 $\hat{\sigma}$ 表示，則

$$\hat{r}_x(t+1) = \hat{k}\hat{\theta} + (1 - \hat{k})\hat{r}_x(t)。$$

二、模擬分析

以臺灣男性 0 歲及 55-59 歲之平均餘命與生存數折扣數列比值為例，說明 BMSDE 的操作方式。首先以 1970-2008 年的折扣數列比值估計參數 k 、 θ 及擴

散係數 σ ，再透過電腦模擬（蒙地卡羅法；Monte Carlo Method）產生亂數，藉由模擬結果可求出折扣數列比值的信賴區間。圖 7 各個圖形中的三條虛線，由上到下分別為 1,000 次模擬的 97.5%、50%（中位數）、2.5% 百分位數，而實線為實際的觀察值。我們發現觀察值大致與中位數接近，而且都落在 95% 信賴區間內，其中可能因為人口數較少，55-59 歲的高年齡層信賴區間比較寬。

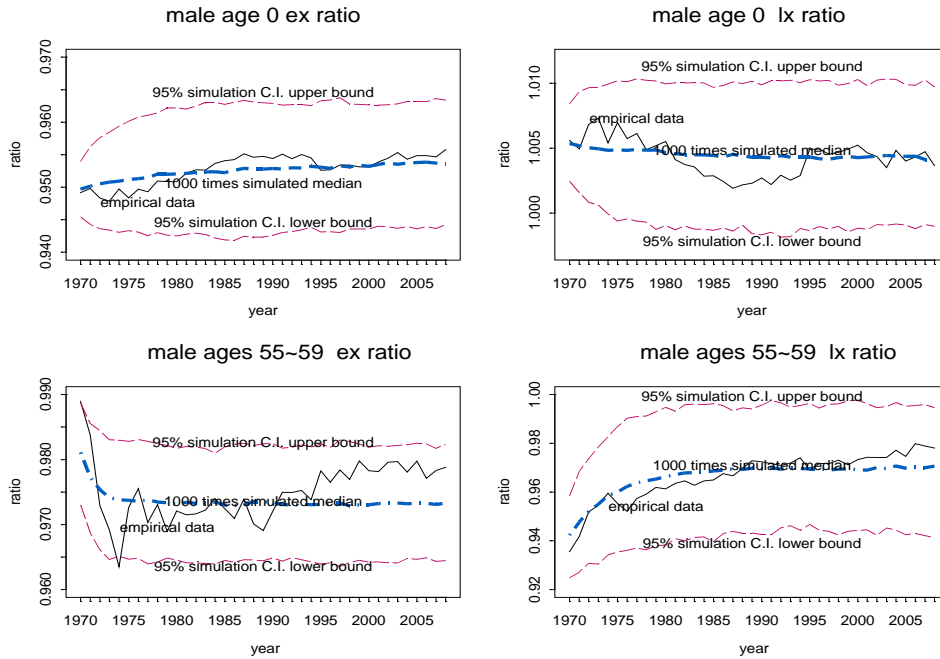


圖 7、臺灣男性 0 歲及 55~59 歲平均餘命與生存數模擬(1970~2008 年)

值得一提的是 0 歲生存數的比值，在前一節發現比值不符合規律性，但使用 BMSDE 卻可以捕捉變化趨勢，而且模擬的平均比值約為 1.005，顯示以生存數描述 0 歲死亡率的時間變化相當穩定。其他年齡的電腦模擬結果大致和圖 7 的類似，在此不再贅述。由電腦模擬結果顯示 BMSDE 確實可用於估計折扣數列比值，接下來再以交叉驗證(Cross-validation)探討 BMSDE 是否可用於預測。

三、交叉驗證

繼續由交叉驗證方法驗證 BMSDE 模型的有效性。使用 HMD 的日本、美國和臺灣生命表，攫取其中的五齡組男、女性的平均餘命與生存數資料，並將資料依年代分成二部分：訓練期(Training Period)與測試期(Testing Period)，訓練期均

為 1970-2000 年，測試期在日本是 2001-2007 年、美國為 2001-2006 年、臺灣為 2001-2008 年。首先以訓練期資料估計 BMSDE 的模型參數，然後給定初始值進行電腦模擬，得出訓練期的估計值、測試期的預測值。圖 8 及圖 9 分別為臺灣幼年（0、1-4、5-9 歲）、高齡層（55-59、60-64、65-69 歲）的平均餘命與生存數電腦模擬結果，實線為實際觀察值、虛線為電腦模擬的結果，日本與美國詳見附錄圖 3~圖 6。

和圖 7 的結果類似，訓練期的 95% 信賴區間涵蓋真實觀察值，而且與模擬的中位數接近；測試期的預測結果也合乎預期，若以測試期模擬的中位數做為未來的預測值，結果也相當不錯。不過或許因為高齡層的人口數較少，測試期的平均餘命和生存數比值，其誤差都比 0 歲大，這從高齡組的信賴區間寬度比幼年組寬一些也可看出。比較性別差異，男性信賴區間比較寬，表示男性死亡率波動較大，女性死亡率的估計相對比較準確。比較三個國家的 BMSDE 模擬結果，發現日本、美國在幼年組比臺灣穩定，高齡組則以日本較為穩定，這和日本平均壽命較高可能有關，顯示未來臺灣和美國的死亡率仍有比較大的變化空間。

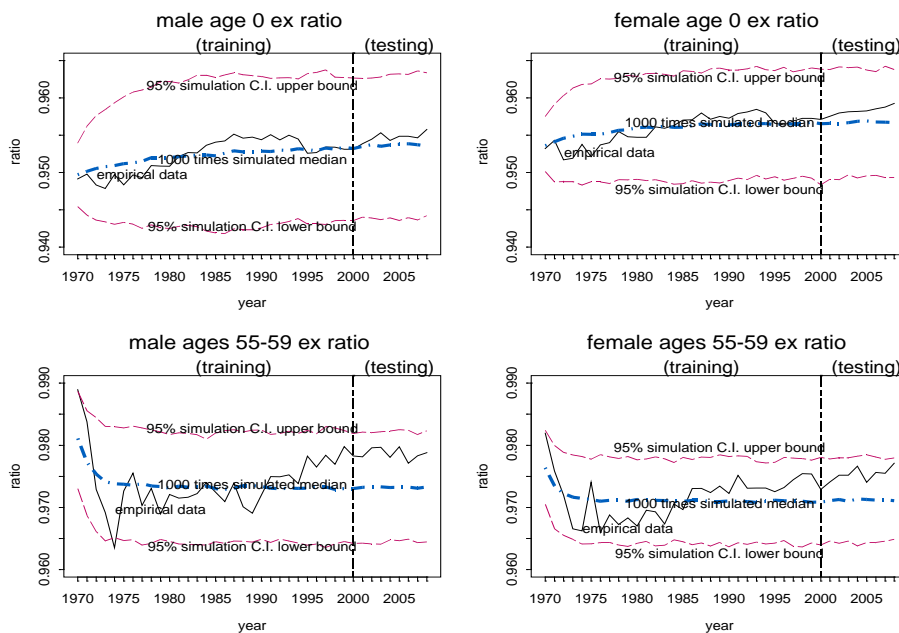


圖 8、臺灣男女兩性 0 歲及 55~59 歲平均餘命模擬(1970~2008 年)

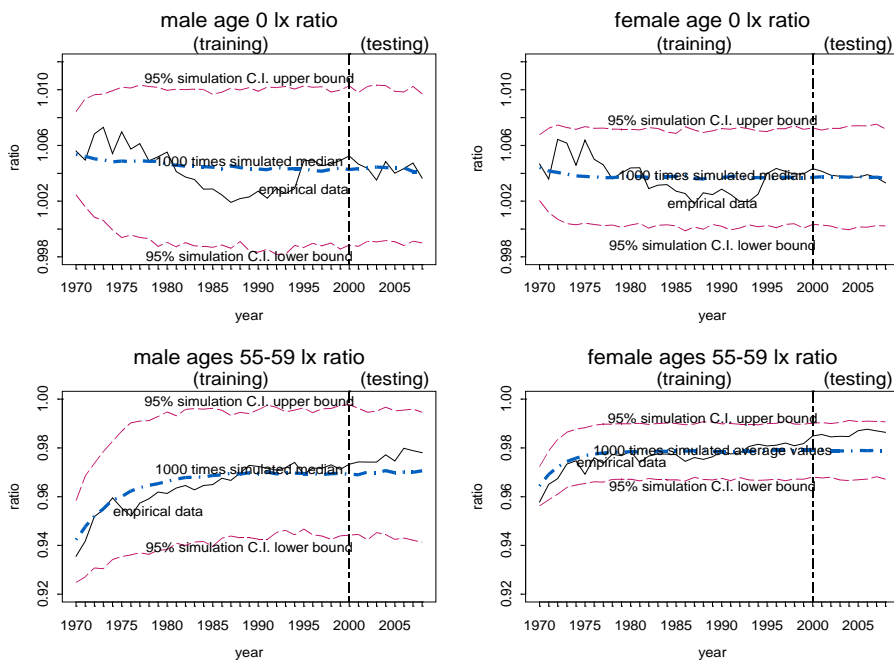


圖 9、臺灣男女兩性 0 歲及 55~59 歲生存數模擬（1970~2008 年）

由上述的模擬分析、交叉驗證，發現 BMSDE 模型確實能捕捉日本、美國、臺灣三國的死亡率變化趨勢，除了美國高齡生存數的預測區間無法涵蓋真實觀察值外，結合折扣數列及 BMSDE 估計法，確實可用於描述及預測死亡率。BMSDE 法可用於描述折扣數列的變化趨勢，預測折扣數列則需要使用其他統計方法，限於篇幅，以下僅提出兩種較為簡易的做法，與常見的 Lee-Carter 模型比較。另外，因為生存數及平均餘命不容易轉化為死亡率，接下來將套用 BMSDE 在死亡率上，但僅以高齡死亡率為探討對象。

四、高齡死亡率推計

折扣數列的死亡率預測可先對死亡率對數值，利用主成分分析(Principal Component Analysis; PCA)縮減變數，作為 BMSDE 模型配適能力改善的分群方法，以台灣男女兩性高齡死亡率(66 歲~85 歲)為研究對象，與 Lee-Carter 方法比較。仿造先前的交叉驗證，將資料分成 1970~2003 年的訓練期與 2004~2008 年的測試期，以訓練期的模型代入測試期，以測試期的平均絕對誤差百分比(MAPE；

Mean Absolute Percentage Error)作為評估標的，結果如表 1。

如同本節一開始的說明，Lee-Carter 模型大致得出各年齡對數死亡率隨時間呈現線性遞降（亦即： $\ln(m_{x,t}) = \alpha_x + \beta_x \cdot t$ ），這個結果可透過 SVD 或 PCA 求得，然而，近年來不少研究都發現，參數 α_x 和 β_x 、甚至是 κ_t 的估計值，會因為使用資料期間而改變，進而影響未來死亡率的預測。本文使用 PCA 預測 BMSDE（兩個主成份），也受到資料期間的影響，直觀上，接近現在的資料應該有較高的權數(Weight)，以此方法與 PCA 結合預測未來的 BMSDE。

除了可透過 PCA 預測 BMSDE，電腦模擬也是近年熱門的預測方法。因為各年度的 BMSDE 間並非獨立，可使用專門處理相依(dependent)資料的區塊拔靴法(Block Bootstrap)或篩網拔靴法(Sieve Bootstrap)，其中區塊拔靴法為無母數方法，根據過去資料特性預測未來，而篩網拔靴法則是加入時間數列的假設，算是一種參數(parametric)方法（參考 Bühlmann, 2002）。區塊拔靴法除了可用於推估未來的死亡率、平均餘命（郭孟坤、余清祥，2007），行政院經濟建設委員會在最近兩次臺灣地區人口推估（2008 年, 2010 年），也使用區塊拔靴法預測生育率及國際遷移（陳政勳、余清祥，2010）。本文也將嘗試以區塊拔靴法預測未來的 BMSDE 數值，區塊拔靴法的設定仿造陳政勳、余清祥，一次抽取區塊長度為五的資料（亦即五個年度），權數則與上述 PCA 類似的線性權數。

表 1、Lee-Carter 與 BMSDE 死亡率 MAPE 比較（臺灣男女兩性）

		Lee-Carter	BMSDE	
			PCA	Block Bootstrap
1970-2003 年 配適誤差	男性	5.95%	7.39%	
	女性	5.47%	7.30%	
2004-2008 年 預測誤差	男性	7.40%	5.69%	5.20%
	女性	10.94%	13.63%	5.21%

表 1 為上述兩種 BMSDE 預測方法、以及 Lee-Carter 模型的死亡率預測比較，其中 BMSDE 預測方法的差異僅在測試期，因此兩種方法在訓練期有相同的

誤差。由表 1 可知，雖然 BMSDE 在男女兩性訓練期的 MAPE 誤差略大於 Lee-Carter 模型，但在測試期的 MAPE 誤差不見得比 Lee-Carter 模型差，區塊拔靴法的預測結果最好，無論男性或女性都明顯優於 Lee-Carter 模型；PCA 和 Lee-Carter 的預測結果和資料有關，男性預測以 PCA 較佳、女性預測以 Lee-Carter 較佳。表 1 考慮的各預測方法，都符合預測誤差在 10%~20%的「良好」標準(Lewis, 1982)，實證上都是可行的方法。

伍、結論與建議

由於醫療科技進步等因素，人口老化是目前世界各國的共同現象。人口老化本身不是問題，但其可能衍生的醫療費用負擔、工作人口減少、老人照護安養等等才是我們需要解決的目標。當一個社會人口老化後，子女奉養壓力沉重，政府勢必承接相當大的照護責任，因此了解未來老人壽命趨勢，精確推估未來人口死亡率，藉以規劃整體規劃人口福利政策，減輕可能影響整體國家與個人的長壽風險，實為當務之急。許多學者對於長壽風險的研究，通常利用一個死亡率指數進行證券化保值，例如生存債券、生存換匯、生存期貨及生存期貨選擇權等，其中死亡率指數的選擇是影響這些商品成功的關鍵，因此能掌握死亡率變動與未來趨勢的死亡率模型相當重要。

由於許多條件的限制，諸如高齡人口死亡率資料的不足，與未來醫療科技發展速度等，增添建立模型的困難、判斷模型是否恰當的困難。有鑑於此，本文引進吃角子老虎問題(Bandit Problem)的折扣數列，驗證常見的生命表函數(例如：死亡率、平均餘命)是否服從折扣數列規律性，並測試這個模型用於高齡死亡率推估的可能性。研究發現許多常用的死亡率模型，例如：Gompertz 法則以及均勻死亡 (Uniform Distribution of Death)假設、定死力(Constant Force)、雙曲線(Hyperbolic)假設等都滿足規律折扣數列的條件。

另外，本文也以臺灣、日本與美國三個國家的實證資料，發現各年齡的平均

餘命及生存數大多符合或接近規律折扣數列；再以布朗運動死亡率隨機微分方程模型(Brownian Motion Mortality Stochastic Differential Equations Model；BMSDE) 規範折扣數列，藉由交叉驗證及電腦模擬，得出三個國家的各年齡平均餘命及生存數的預測值，都在 BMSDE 的 95%信賴區間內。本文也以預測臺灣地區 66-85 歲男女兩性高齡死亡率為目標，與常見的 Lee-Carter 模型比較，發現結合 BMSDE 與折扣數列，以區塊拔靴法得出的死亡率預測誤差優於 Lee-Carter 模型。綜合這些結果，顯示規律折扣數列在實證上可行，作為高齡死亡率的另一種選擇。因為折扣數列模型在臺灣高齡死亡率的預測誤差較小，可用於政府推估未來高齡人口、人壽保險業計算年金險等商品，降低死亡率預測（以及精算費率）的誤差，以因應人口老化及長壽風險的相關問題。

另外本文研究也發現以下三項結果：

第一項是平均餘命、生存數、死亡數與死亡率的折扣數列，在理論推導與經驗資料分析方面大多具有規律性，並且隨著時間增加，比值有趨近定值的現象，因此折扣數列比值的規律性，應可以做為死亡率模型選取的判斷準則。

第二項是由經驗資料分析中，發現各年齡組生存數比值會隨著年齡增加呈現指數遞減，並且有穩定收斂的情形，因此可計算同一年齡生存數比值的中間數，再以最小平方法估計 BMSDE 的參數值，作為估計或改善估計各國高齡人口死亡率的參考。

第三項是若死亡率資料有遺漏值，且插補方法不適合時，也可使用折扣數列比值估計這些遺漏值。因為 BMSDE 掌握了每個年齡隨時間的變動，可以利用來推估遺漏值，或是透過時間序列方法預測未來死亡率與平均餘命。

另外，本文第四節的死亡率預測比較，限於篇幅，只以臺灣男女兩性為探討對象，未來將繼續代入 HMD 中的其他國家，以確定規律折扣模型是否穩定、且有準確的死亡率預測。初步測試發現日本和美國的死亡率預測，結果和表 1 類似，區塊拔靴法的死亡率的誤差較小，明顯優於 PCA 的預測，因此我們推薦以區塊拔靴法結合 BMSDE；但由於這些方法大多透過電腦模擬，預期未來推廣時，

操作人員需要更多相關訓練及知識。除了電腦模擬的區塊拔靴法外，近來多變量分析中的函數主成份分析 (Functional PCA) 的相關討論較多(Hyndman and Ullah, 2007; Ramsay and Silverman, 2005)，將探討使用 FPCA 的可能性。

死亡率模型中加入時間改善因子，除了第一節中介紹的模型外，也有使用和本文想法類似的方法。例如：林麗芬、張燕明(2005)將凌波理論應用於死亡率改善幅度，英國精算學會(1999)的連續死亡率調查報告第 17 號，對英國 60 歲以上公務人員的縮減因子(Reduction Factor)死亡率模型，均是年齡死亡率比值的模型，但這些模型大多假設死亡率改善幅度為定值。本文提出折扣數列比值描述死亡率的改善，雖然也是改善也與時間有關，但改善趨勢不見得是時間的定值。

有鑑於許多實證研究都發現死亡率改善未必是定值，近年來死亡率模型的研究上經常加入世代的影響，本文使用的布朗運動方法也有必要考量世代及相鄰年代與相鄰年齡的影響，應有助於掌握死亡率的隨機變化。另一個可能方向是將年齡與時間視為二維座標，以空間統計(Spatial Statistics)方式描述死亡率的變化，未來我們將引進空間統計分析死亡率。此外，本文發現折扣數列比值有逐漸收斂的傾向，可以採用類似終壽區間成數(Fraction of the Last Age Interval of Life)的想法，以定常人口的角度估計死亡率，並作為建構生命表之用；將本文使用的 BMSDE 與其他死亡率模型（例如：Lee-Carter）結合，以改進死亡率的配適，也是未來可能的研究方向。

謝誌

本文作者感謝匿名評審的寶貴建議，更感謝國科會對本文在研究進行與撰寫時的補助（補助計畫編號NSC 99-2410-H-004 -064 -MY2）。

參考文獻

中文部分

- 行政院經濟建設委員會（2008）中華民國臺灣97年至145年人口推計。
- 行政院經濟建設委員會（2010）2010年至2060年臺灣人口推計。
- 林麗芬、強燕明（2005）凌波理論於死亡率改善幅度之預測，逢甲大學統計與精算研究所碩士論文。
- 陳政勳、余清祥（2010）“小區域人口推估研究：臺北市、雲嘉兩縣、澎湖縣的實證分析”。人口學刊，41:153-183。
- 郭孟坤、余清祥（2007）“電腦模擬、隨機方法與人口推估的實證研究”。人口學刊，36: 67-98。

英文部分

- Berry, D.A. and Fristedt, B. 1979. “Bernoulli One-Armed Bandits--Arbitrary Discount Sequences.” *The Annals of Statistics* 7(5) : 1086-1105.
- Berry, D.A. and Fristedt, B. 1985. *Bandit Problems*. Chapman & Hall.
- Bühlmann, P. 2002. “Bootstraps for Time Series.” *Statistical Science* 17(1): 52-72.
- Cairns, A.J.G., Blake, D., Dowd, K. 2006a. “Pricing death: Framework for the valuation and securitization of mortality risk.” *ASTIN Bulletin* 36: 79–120.
- Cairns, A.J.G., Blake, D., Dowd, K. 2006b. “A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration.” *Journal of Risk and Insurance* 73(4): 687-718.
- Coale, A.J. and Kisker, E.E. 1990. “Defects in data on old-age mortality in the United States: new procedures for calculating mortality schedules and life tables at the highest ages.” *Asian and Pacific Population Forum* 4(1): 1-31.
- Continuous Mortality Investigation Bureau (CMIB) 1999. “*Mortality Improvement Model for use with the 92'Series of Mortality Tables*.” Continuous Mortality Investigation Reports No 17.
- Delwarde, A., Denuit, M., and Eilers, P. 2007a. “Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: A penalized log-likelihood

- approach.” *Statistical Modeling* 7: 29-48.
- Chen, H. and Cox, S. 2009. “Modeling Mortality with Jumps: Applications to Mortality Securitization.” *Journal of Risk and Insurance* 76(3):727.
- Heligman, L.M.A. and Pollard, J.H. 1980. “The age pattern of mortality.” *Journal of the Institute of Actuaries* 107(1): 49-82.
- Hyndman, R. J. and M. S. Ullah. 2007. “Robust Forecasting of Mortality and Fertility Rates: A Functional Data Approach.” *Computational Statistics & Data Analysis* 51(10): 4942-4956.
- Lee, R.D. and Carter, L. 1992. “Modeling and forecasting the time series of U.S. mortality.” *Journal of the American Statistical Association* 87: 659-675.
- Lewis, E. B. 1982. Control of body segment differentiation in *Drosophila* by the bithorax gene complex, *Embryonic Development, Part A: Genetics Aspects*, Edited by Burger, M. M. and R. Weber. Alan R. Liss, New York, 269-288.
- Ramsay, J. O. and B.W. Silverman. 2005. *Functional Data Analysis*. 2nd ed. New York: Springer-Verlag.
- Renshaw, A.E. and Haberman, S. 2006. “A cohort-based extension to the Lee-Carter model for mortality reduction factors.” *Insurance: Mathematics and Economics, Elsevier* 38(3):556-570.
- Yang, S.S., Yue, C.J., and Huang, H.C. 2010. “Modeling Longevity Risks using a Principal Component Approach: A Comparison with Existing Stochastic Mortality Models.” *Insurance Mathematics and Economics, Elsevier*.46(1) : 254-270
- Yue, C. J. (1999), “Generalized Two-stage Bandit Problem”, *Communications in Statistics: Theory and Methods*, vol. 28(9), 2261-2276.

附錄

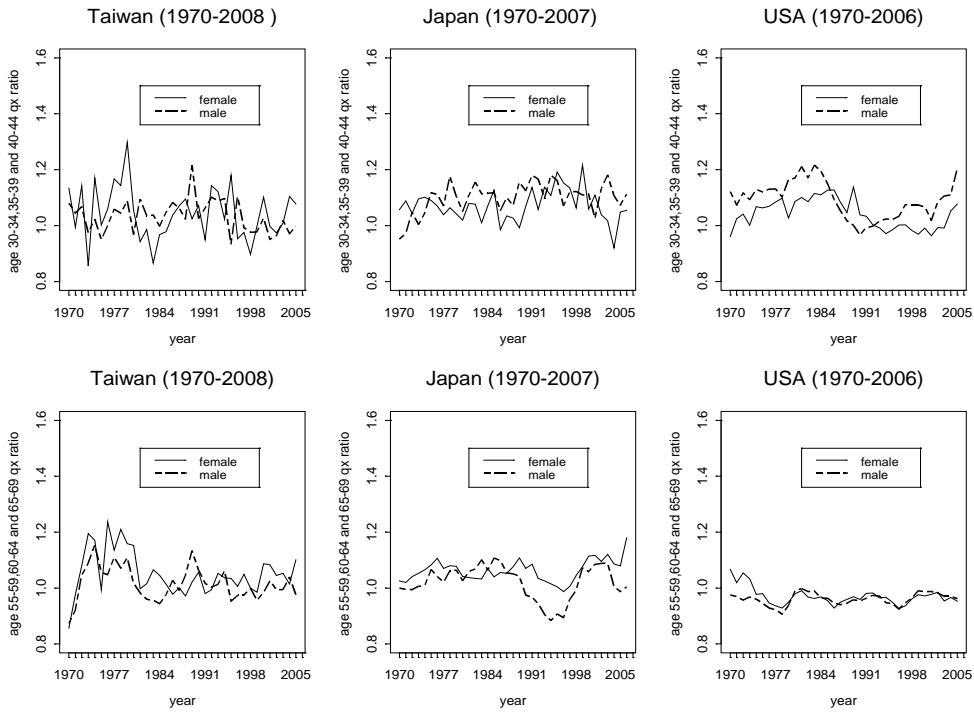


圖 1、青壯年與老年兩個年齡層死亡率折扣數列比值

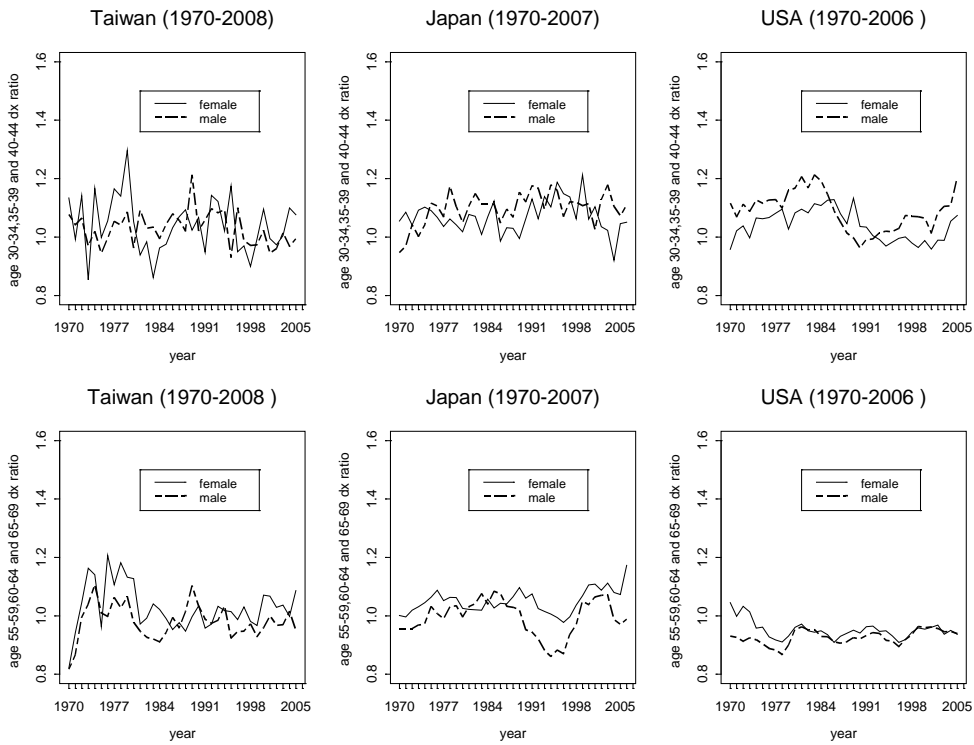


圖 2、青壯年與老年兩個年齡層死亡數折扣數列比值

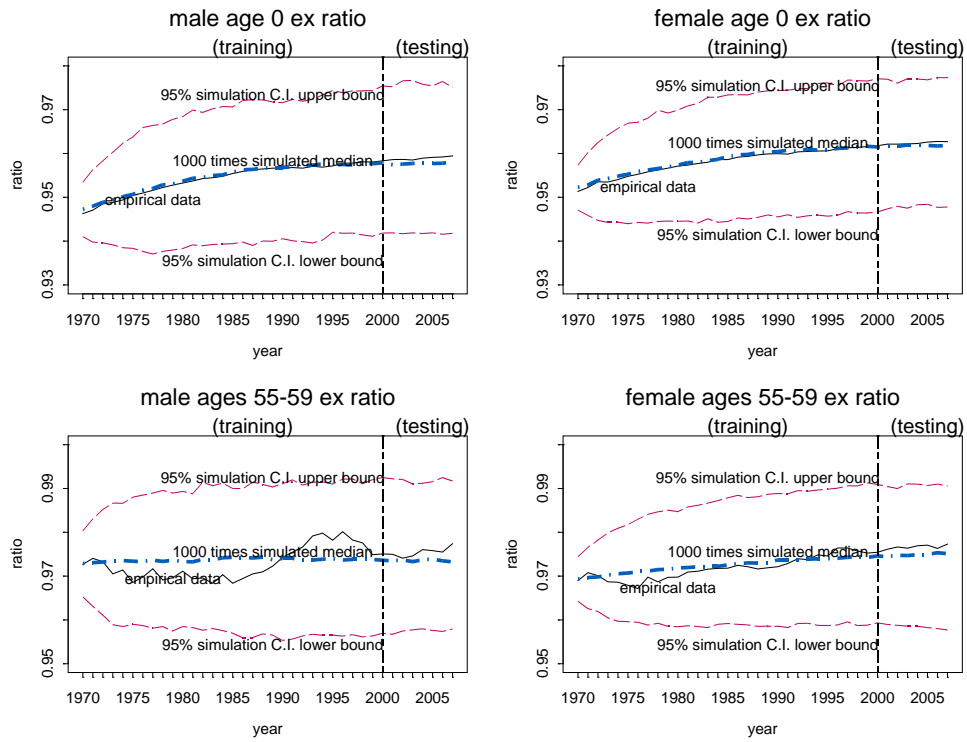


圖 3、日本男女兩性 0 歲及 55~59 歲平均餘命模擬(1970~2007 年)

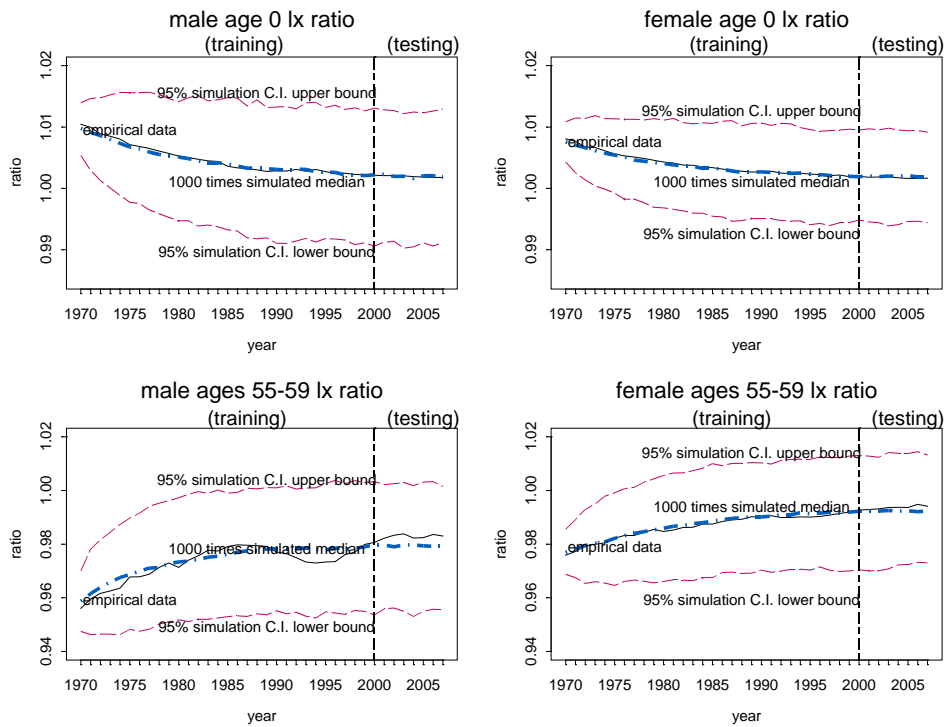


圖 4、日本男女兩性 0 歲及 55~59 歲生存數模擬(1970~2007 年)

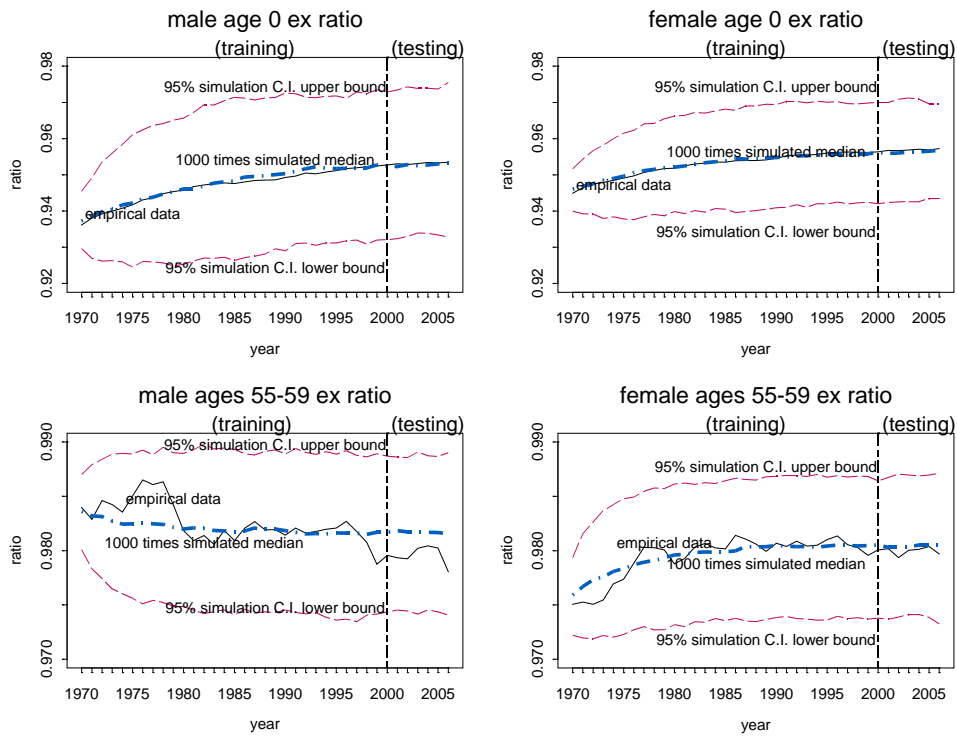


圖 5、美國男女兩性 0 歲及 55~59 歲平均餘命模擬(1970~2006 年)

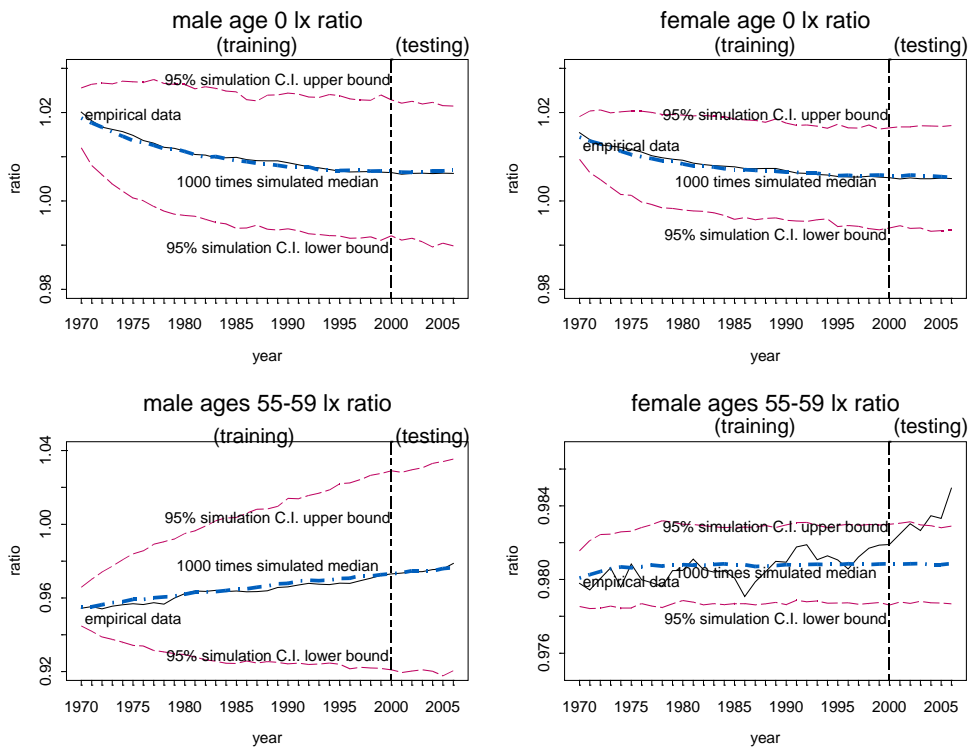


圖 6、美國男女兩性 0 歲及 55~59 歲生存數模擬(1970~2006 年)