修勻學(Graduation)－
生命表的建構與相關考量

教師：余清祥教授

日期：2024年9月25日

下載：http://csyue.nccu.edu.tw

生命表(Life Table)

- 生命表(Life table)為計算生命函數（包括死亡率、平均壽命、……）的基礎表，根據一年或數年觀察時間之統計資料編算而成，通常分為國民生命表(Population Table)及經驗生命表(Experience Table) 。
- 在台灣地區國民生命表由政府部門修訂，經驗生命表由保險公司或其公會編纂。

生命表的歷史發展

- 生命表可追溯自16世紀後期，歐洲各教區(Parish)紀錄信眾的人口狀況。

→Graunt於1662年公佈倫敦居民的死亡率(*Observations on the London Bills of Mortality*)；

→哈雷於1693年完成第一個生命表(Breslau)；

→數學家Deparcieux (1746)、Bernoulli (1766)、Lambert (1772)分別對於平均餘命、μ(x)、S(x)有所貢獻。

生命表的歷史發展(續)

- In 1725 de Moivre first proposed a piecewise linear survivor function $S(x) - S(x+t) = a\ t$.
- In 1772 Lambert first expresses the gradual 'exhaustion of man's power' in a mathematical formula.
- In 1798 Malthus predicted exponential increase of human population with consequent misery and famine unless family size was regulated. The model was later adjusted by Verhulst in 1838, to limit the population growth predictions.
- In 1825 Gompertz proposed a survivor function which results in an exponentially rising mortality intensity. This model was later extended by Makeham in 1860 to allow for young mortality.
- In 1872 Thiele develops the first complex model that attempts to accurately model $\mu(x)$ over the whole life span.

生命表的應用

Process	*State studied*	*State entered when...*	*State left when...*	*Vertical dimension of the table*
Mortality	Being alive	Born	Die	Duration of life (age)
Nuptiality (first marriage)	Being unmarried	Born	Marry	Duration of single life (age)
Migration from place of birth	Living in place of birth	Born	Move to another place	Duration of residence (age)
Entering the labor force	Having never worked	Born	First enter labor force	Duration of life (age)
Becoming a mother	Having no births	Born	Have first birth	Duration of life (age)
Subsequent childbearing	Not having an additional birth	Have a birth	Have an additional birth	Duration since having a birth
Marital survival	Being in intact marriage	Marry	Marriage ends	Duration of marriage
Unemployment spells	Being unemployed	Become unemployed	Leave state of unemployment	Duration of unemployment
Incarceration	Being in jail	Enter jail	Leave jail	Duration of incarceration

[a] All of these processes can also be conceived as multiple decrement processes. Here we ignore other risks of leaving the state of interest.

國民生命表的編算流程

- 建構生命表需要各年齡人數與死亡人數，若非擁有完整人口記錄（台灣有戶籍登記），每十年一次的戶口普查才有各年齡人數。

→另外，死亡人數來自生命統計。

- 完全生命表一般以普查年為中心點，蒐集前後一年共三年的死亡資料編製而成。

→台灣亦是如此，每十年編算一次完全生命表，每年一次簡易生命表。

1969年臺灣地區簡易生命表

男性

年齡組	死亡機率	生存數	死亡數	定常人口		平均餘命
X ~ (X+n)	qx	lx	dx	Lx	Tx	ex
0	0.01844	100000	1844	98626	6633741	66.34
1 — 4	0.01328	98156	1304	389102	6535115	66.58
5 — 9	0.00404	96852	392	483155	6146012	63.46
10 — 14	0.00325	96461	313	481578	5662857	58.71
15 — 19	0.00612	96147	588	479386	5181279	53.89
20 — 24	0.00865	95559	827	475819	4701893	49.20
25 — 29	0.01078	94733	1022	471168	4226075	44.61
30 — 34	0.01245	93711	1166	465735	3754907	40.07
35 — 39	0.01637	92545	1515	459099	3289173	35.54
40 — 44	0.02221	91030	2022	450372	2830074	31.09
45 — 49	0.03401	89008	3027	438022	2379702	26.74
50 — 54	0.05278	85981	4538	419084	1941680	22.58
55 — 59	0.07532	81443	6134	392915	1522596	18.70
60 — 64	0.12753	75309	9604	354002	1129680	15.00
65 — 69	0.19535	65705	12835	297545	775678	11.81
70 — 74	0.28679	52870	15163	227208	478133	9.04
75 — 79	0.42151	37707	15894	148426	250925	6.65
80 — 84	0.59018	21813	12874	74866	102499	4.70
85+	1.00000	8940	8940	27633	27633	3.09

第十次(民國108~110年)臺灣地區國民生命表

民國 108-110 年 男性

年齡 X	生存數 l_x	死亡數 d_x	生存機率 p_x	死亡機率 q_x	定常人口 L_x	定常人口 T_x	平均餘命 x
日 DAY							
0	100000	211	0.99790	0.00211	1916	7741155	77.41
7	99790	26	0.99974	0.00026	1914	7739239	77.56
14	99764	19	0.99981	0.00019	1913	7737325	77.56
21	99745	11	0.99989	0.00011	1913	7735412	77.55
28	99734	36	0.99964	0.00036	8742	7733499	77.54
月 MONTH							
2	99699	18	0.99982	0.00018	8194	7724757	77.48
3	99681	51	0.99949	0.00051	24573	7716563	77.41
6	99630	43	0.99957	0.00043	50486	7691991	77.21
年 YEAR							
0	100000	416	0.99584	0.00416	99792	7741155	77.41
1	99584	136	0.99864	0.00136	99516	7641363	76.73
2	99448	71	0.99928	0.00072	99413	7541846	75.84
3	99377	29	0.99971	0.00029	99363	7442434	74.89
4	99348	16	0.99984	0.00016	99340	7343071	73.91
5	99332	14	0.99986	0.00014	99325	7243731	72.92
6	99318	12	0.99988	0.00012	99313	7144406	71.93
7	99307	10	0.99989	0.00011	99302	7045093	70.94
8	99296	10	0.99990	0.00010	99291	6945792	69.95
9	99286	11	0.99989	0.00011	99281	6846500	68.96

臺灣壽險業經驗生命表

1,000qx

年齡 Age	臺灣壽險業第六回經驗生命表 2021 Taiwan Standard Ordinary Experience Mortality Table (2013-2017)				臺灣壽險業第五回經驗生命表 2011 Taiwan Standard Ordinary Experience Mortality Table (2004-2008)			
	男 性 Male		女 性 Female		男 性 Male		女 性 Female	
	死亡率 qx	平均餘命 $\mathring{e}_x$	死亡率 qx	平均餘命 $\mathring{e}_x$	死亡率 qx	平均餘命 $\mathring{e}_x$	死亡率 qx	平均餘命 $\mathring{e}_x$
0	0.320	81.11	0.250	86.40	0.522	77.14	0.389	83.20
1	0.189	80.13	0.145	85.66	0.384	76.18	0.304	82.23
2	0.163	79.15	0.124	84.67	0.277	75.21	0.218	81.25
3	0.140	78.16	0.105	83.68	0.215	74.23	0.183	80.27
4	0.125	77.17	0.093	82.69	0.181	73.25	0.158	79.28
5	0.114	76.18	0.083	81.70	0.166	72.26	0.138	78.30
6	0.111	75.19	0.080	80.71	0.149	71.27	0.121	77.31
7	0.112	74.20	0.078	79.71	0.139	70.29	0.110	76.32
8	0.114	73.21	0.076	78.72	0.134	69.30	0.103	75.33
9	0.119	72.21	0.075	77.72	0.133	68.30	0.101	74.33
10	0.122	71.22	0.070	76.73	0.129	67.31	0.103	73.34
11	0.137	70.23	0.072	75.74	0.131	66.32	0.110	72.35
12	0.155	69.24	0.077	74.74	0.153	65.33	0.123	71.36
13	0.181	68.25	0.085	73.75	0.196	64.34	0.141	70.36
14	0.227	67.26	0.097	72.75	0.255	63.35	0.159	69.37
15	0.296	66.28	0.130	71.76	0.344	62.37	0.181	68.39
16	0.339	65.30	0.144	70.77	0.455	61.39	0.206	67.40
17	0.378	64.32	0.157	69.78	0.540	60.42	0.232	66.41
18	0.410	63.34	0.169	68.79	0.584	59.45	0.243	65.43
19	0.435	62.37	0.181	67.80	0.607	58.48	0.249	64.44
20	0.432	61.40	0.178	66.81	0.624	57.52	0.253	63.46
21	0.447	60.42	0.187	65.83	0.641	56.56	0.259	62.47
22	0.459	59.45	0.196	64.84	0.668	55.59	0.273	61.49
23	0.466	58.48	0.203	63.85	0.710	54.63	0.295	60.51

- **簡易生命表(Abridge Life Tables)**

  生命表有時因為需要，必須每年編算一次，但因為

  →資料不足、缺乏或是品質無法保證

  →使用者不需要太過詳細的資料，編算時分為 0、1-4歲、5-9、10-14、…(五齡組，台灣最高年齡組為85＋)。

- 簡易生命表基本上由中央死亡率出發，再轉成死亡率：

$$ {}_nm_x = \frac{{}_nd_x}{{}_nL_x} \quad \Rightarrow \quad {}_nq_x $$

經驗生命表的編算流程

- 編算生命表需要先蒐集經驗理賠資料，包括壽險業各類型商品的投保人數（或暴露數；Exposure）、死亡人數。先由理賠資料得出原始死亡率，再使用修勻方法得出較為平滑的死亡率，最後得出生命表。

→保險業的生命表可分成死亡險生命表(e.g.第五回經驗生命表)、生存險生命表(e.g.第二回年金生命表)等。

我國壽險業經驗生命表編製歷程

名稱	資料時間	編製時間	編製單位	主管機關核准時間及函釋
臺灣壽險業生命表	民國58～61年	民國62～63年	臺北市人壽保險商業同業公會	民國64年2月5日台財錢字第11200號
臺灣壽險業第二回生命表	民國66～70年	民國70～72年	臺北市人壽保險商業同業公會	民國73年12月18日台財融字第24549號
臺灣壽險業第三回生命表	民國71～75年	民國77～78年	臺北市人壽保險商業同業公會	民國78年6月19日台財融字第780163364號
臺灣壽險業第四回生命表	民國84～88年	民國91年	中華民國人壽保險商業同業公會	民國91年12月27日台財保字第0910074199號
臺灣壽險業第五回生命表	民國95～99年	民國100年	保險發展中心	民國101年1月10日金管保財字第10102500605號
臺灣壽險業第六回生命表	民國102～106	民國108年	保險發展中心	民國110年3月29日金管保財字第11004909551號

- 中央死亡率與死亡率

$$ {}_nM_x = \frac{{}_nD_x}{{}_nL_X} \quad and \quad {}_nq_x = \frac{{}_nD_x}{l_x} $$

→唯一的區別在於分母。

→但通常只能獲得年中人數、年底人數，在定常人口及均匀死亡(UDD)假設下，各年齡別的年中人數可視為 L_x，通常先求出中央死亡率及下式求出死亡率(UDD)：

$$ q_x = \frac{m_x}{1 + m_x / 2}. $$

- 因為單一年度的單齡死亡人數通常不多，尤其當該年齡層的人數及死亡率較低時，編算死亡率可能會有較大的震盪。

$$D_x \sim B(N_x, q_x) \Rightarrow Var(\hat{q}_x) = Var(D_x / N_x) = \frac{q_x(1-q_x)}{N_x}$$

- 可能解決方法： ${}_5M_x \Rightarrow M_x$

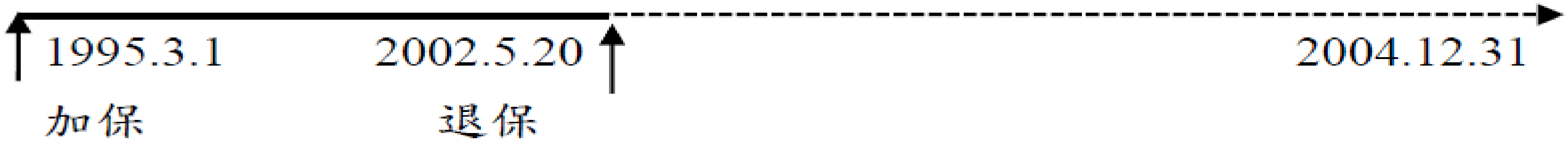
→合併幾個年齡(一般為五歲組)，再以內插(修勻)法找出單一年齡死亡率。

→考慮連續幾年(一般為三年或五年等奇數年)的死亡資料。

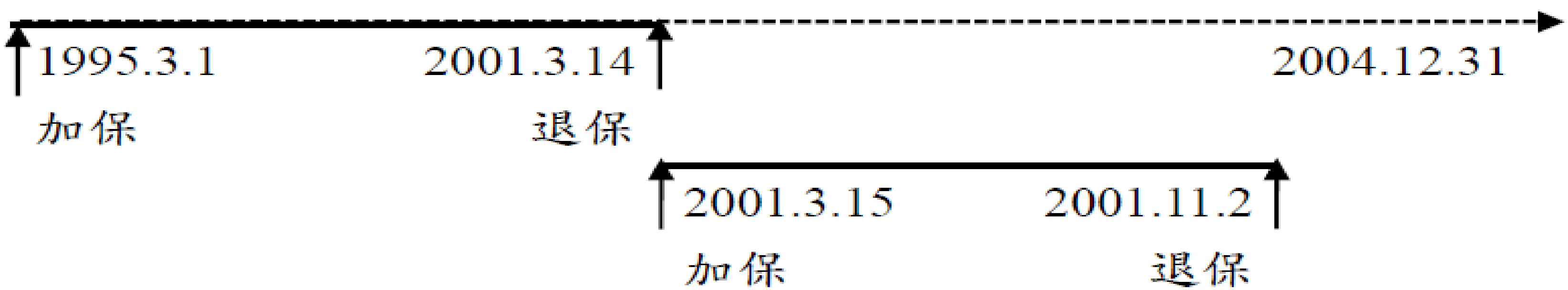
關於暴露數的計算

- 暴露數(Exposure)又稱暴險數，指的是暴露於風險的時間，亦即進入、離開某個狀態的總時間。如下圖：

圖示 1：



圖示 2：



暴露數的計算（續）

- 被觀察者進入研究時間未必為整數年齡，只能考量符合要求的部分。

→1992/7/1出生者，在計算2013年30歲暴露數時，僅有2013/1/1~2013/6/30符合條件，亦即只有30歲的後半歲滿足條件；

→1993/5/1出生者，符合2013年30歲的暴露數，僅有2013/5/1~2013/12/31，或是滿30歲後的前半歲滿足條件。

註：截斷與設限(Truncation & Censor)。

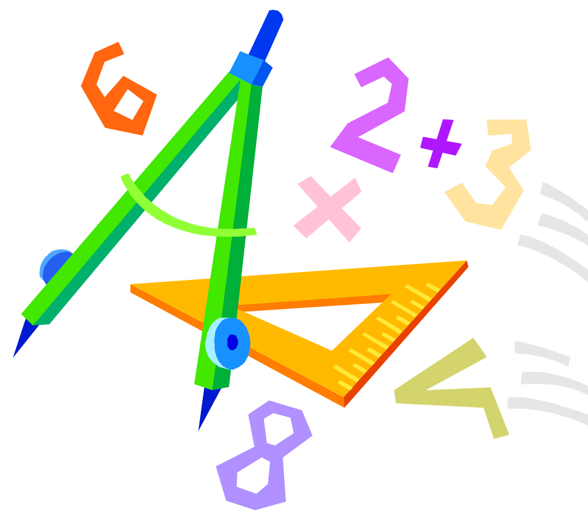
暴露數的近似計算

- 加總每人的暴露數，即可得出各年齡暴露數。不過，有時為了簡化計算，會採用類似「360/30」，亦即假設每月有30天、一年有360天。
- 另外，國民生命表編算通常不會逐一考量個體的暴露數，而是以「各年齡年中人口數」代替暴露數，或是年初人數、年底人數的平均數。

生命表基本函數

- 生存數 l_x
- 死亡數 $_nd_x$
- 死亡率 q_x 及生存活率 p_x
- 定常人口 L_x 及 T_x
- 其他基本函數

→平均餘命 $\mathring{e}_x$



- 死亡分配的特例：

→均勻死亡(Uniform distribution of death;UDD)

$$l_{x+t} = (1-t)\,l_x + t \cdot l_{x+1} \quad , 0 \le t \le 1$$

表示各年齡生存人數隨年齡直線下降。

→在此假設下，定常人口等於

$$L_x = l_{x+1} + \frac{1}{2}d_x = l_x - \frac{1}{2}d_x = \frac{1}{2}\left(l_x + l_{x+1}\right).$$

→均勻死亡假設對死亡率而言，

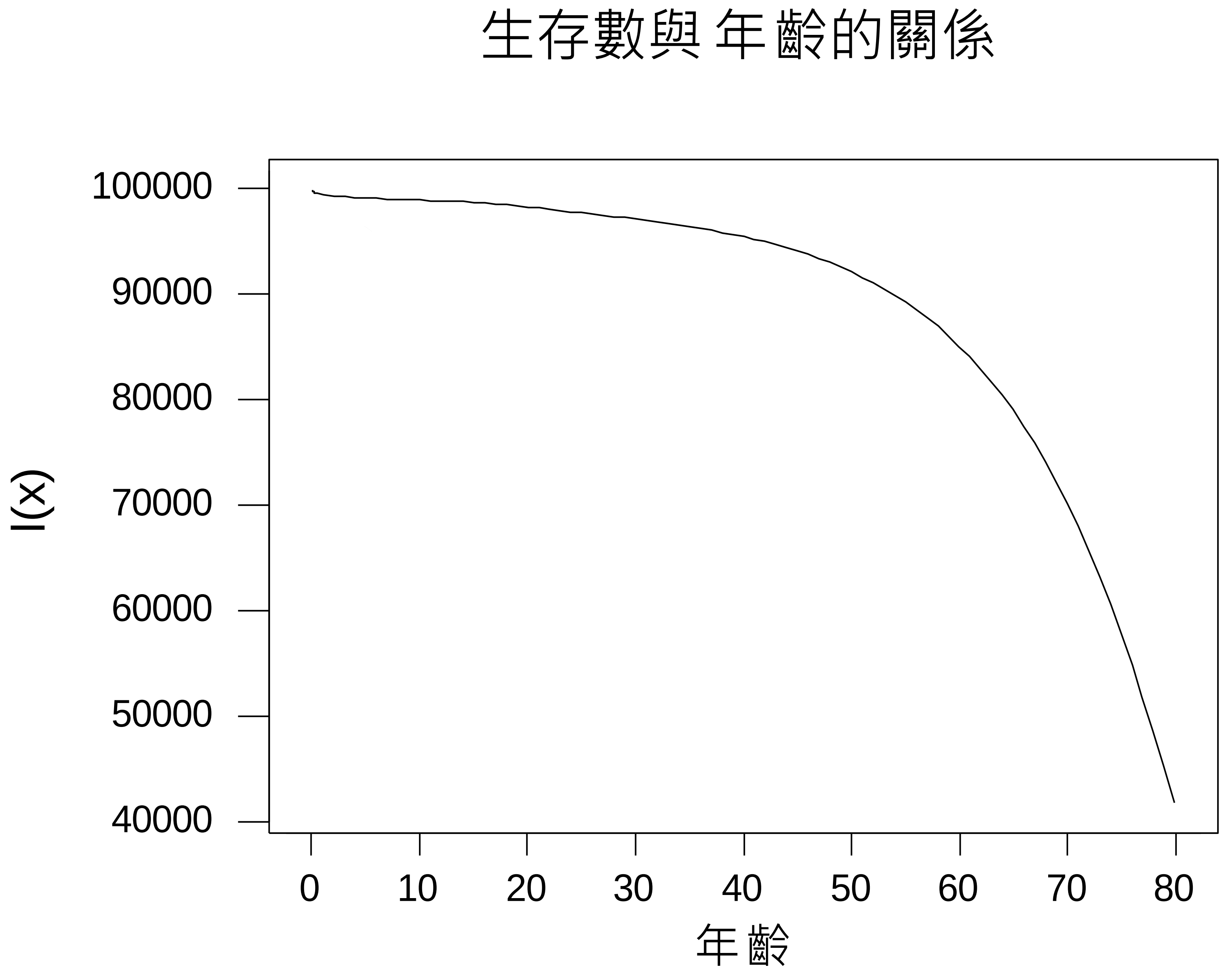
$${}_t q_x = t \cdot q_x$$

- 平均餘命(Expectation of Life)

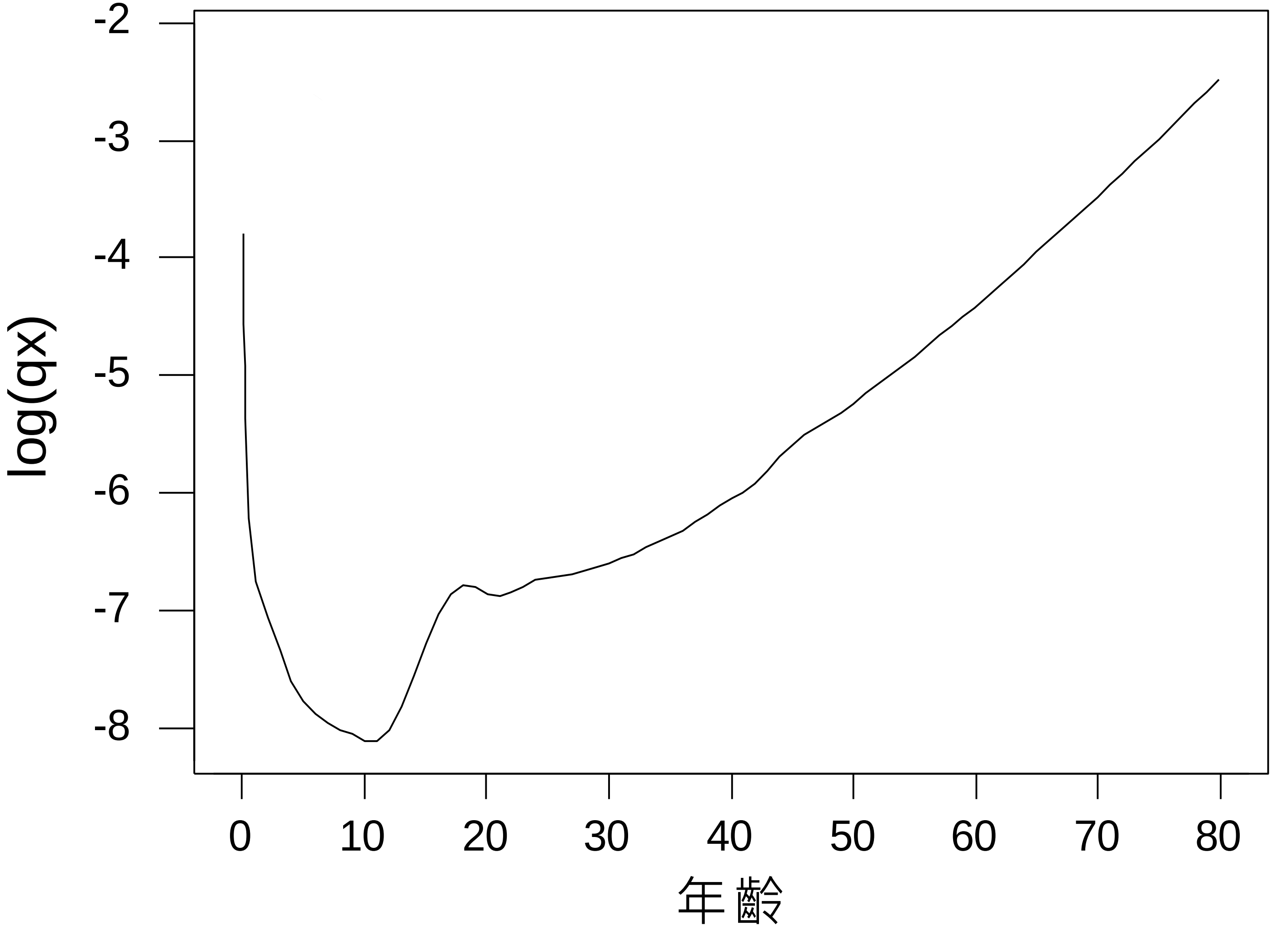
→平均餘命 $\overset{o}{e}_x$ 為生命表中已存活至 x 歲的人，未來預期可存活的年數；在 $x = 0$ 時， $\overset{o}{e}_x$ 代表的即是平均壽命。

→定常人口(Stationary Population)的假設下：

$$\overset{o}{e}_x = \frac{T_x}{l_x}$$


生存數與年齡的關係
l(x)
100000
90000
80000
70000
60000
50000
40000
0
10
20
30
40
50
60
70
80
年齡

log(死亡率)與年齡的關係



死力 μ_x
(瞬間死亡率Force of Mortality)

- 連續變數可用機率分配函數描述其瞬間變化的程度，存活時間為連續變數時可考慮死亡率的瞬間變化：

$$\mathrm{P}(\,\mathrm{x} < \mathrm{X} \le \mathrm{x} + \Delta t \mid X > x) = \mathrm{P}(\,0 < \mathrm{T(x)} \le \Delta t)$$

$$= \frac{S(x) - S(x + \Delta t)}{S(x)}$$

$$= -\frac{S'(x)}{S(x)} \cdot \Delta t$$

因此

$$\mu_x = \lim_{\Delta t \to 0}\left(\frac{{}_{\Delta t}q_x}{\Delta t}\right) = -\frac{S'(x)}{S(x)} = -\frac{1}{S(x)} \cdot \frac{dS(x)}{dx}.$$

將 dx 移至等號左側，

$$\int \mu_x dx = \int -\frac{1}{S(x)} dS(x) = -\log(S(x))$$

$$\Leftrightarrow \quad S(x) = \exp(-\int \mu_x dx).$$

換言之，

$${}_n p_x = \exp(-\int_0^n \mu_{x+t} dt).$$

將存活時間以隨機變數的形式表達，因為

$$ {}_tq_x = 1 - {}_tp_x = P(T(x) \le t), $$

可視為CDF，對時間 t 微分可得PDF：

$$ \frac{d}{dt}({}_tq_x) = -\frac{d}{dt}({}_tp_x) = -\frac{d}{dt}\left(\frac{S(x+t)}{S(x)}\right) $$

$$ = -\frac{1}{S(x)}(-\mu_{x+t} \cdot S(X+t)) = {}_tp_x\, \mu_{x+t}. $$

換言之，${}_tp_x\,\mu_{x+t}$ 可視為存活時間的PDF。

- 代入PDF，死亡機率也可寫成

$$ {}_nq_x = \int_0^n \mu_{x+t}\ p_{x+t}\, dt. $$

- 以連續變數的型態，定常人口可表為

$$ L_x = \int_0^1 l_{x+t}dt = l_x \int_0^1 {}_tp_x dt. $$

x 歲的人之平均餘命等於

$$ \overset{o}{e}_x = \frac{T_x}{l_x} = \frac{l_x \int_0^\infty {}_tp_x\, dt}{l_x} = \int_0^\infty {}_tp_x\, dt. $$

- 將x歲的人之存活時間以 T表示，則

$$E(T)=\overset{o}{e}_x$$

$$Var(T)=\int_0^{\infty} t^2 \mu_{x+t}\ {}_tp_x\, dt-(\overset{\circ}{e}_x)^2$$

$$=2\int_0^{\infty} t\ {}_tp_x\, dt-(\overset{\circ}{e}_x)^2.$$

- E(T) 及 Var(T) 不見得存在。

→例如：$S(x) = (1+x)^{-1}$

或 $S(x) = (1+x)^{-2}$ 。

非整數年齡(Fractional Ages)的假設

- 生命表只有整數年齡的機率分配，因此必須加上非整數年齡的函數假設。
- 常見非整數年齡的假設有三種，給定整數年齡x與$x+1$數值，以內插法求x與$x+1$間數值：

→線性內插(Linear Interpolation)

→指數內插(Exponential Interpolation)

→調和內插(Harmonic Interpolation)

- 線性內插：$(0 \leq t \leq 1)$

  存活函數滿足

$$S(x+t)=(1-t)\cdot S(x)+t\cdot S(x+1)$$

  也就是常見的算術平均，在人口統計學則稱為均勻的死亡分配(U.D.D.)

- 在這個假設下較重要的特性是${}_tp_x$ 為 t 的線性函數：

$${}_tq_x = t\cdot q_x.$$

- 指數內插： $(0 \le t \le 1)$

  存活函數滿足

$$S(x+t) = S(x)^{1-t} \times S(x+1)^{t}$$

  也就是常見的幾何平均，因為在這個假設下死力$\mu_x = \mu$，因死又稱為定死力(Constant Force of Mortality)。

- 在這個假設下較重要的特性是：

$${}_t p_x = (p_x)^t.$$

$$\log(S(x+t)) = (1-t)\log(S(x)) + t\log(S(x+1)).$$

- 調和內插：$(0 \le t \le 1)$

  存活函數滿足

$$\frac{1}{S(x+t)} = \frac{1-t}{S(x)} + \frac{t}{S(x+1)},$$

  又稱為雙曲線(Hyperbolic)假設。

- 在這個假設下較重要的特性是：

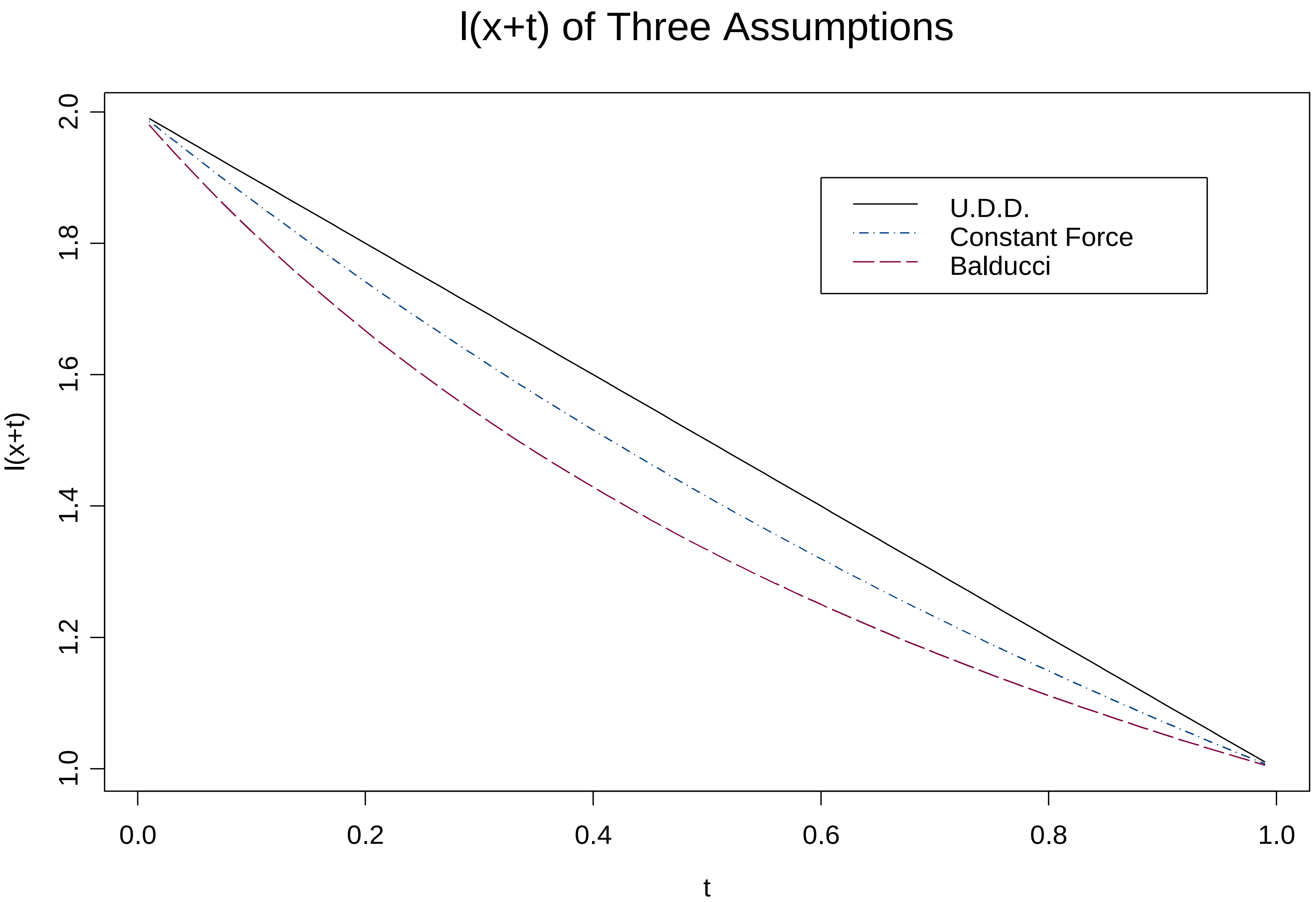
$${}_{1-t}q_{x+t} = (1-t)q_x.$$

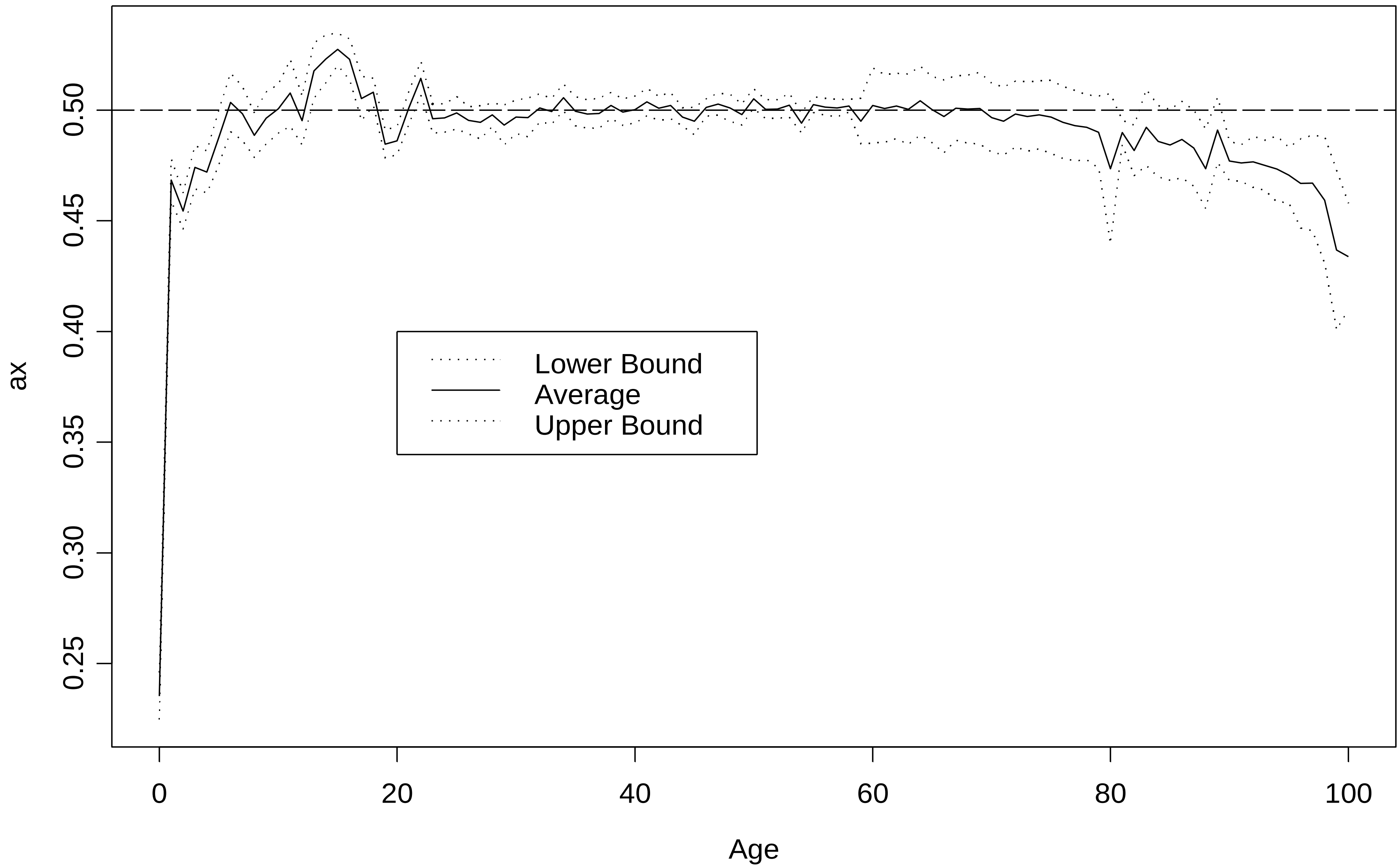
- 終壽區間成數(Fraction of the Last Age Interval of Life)：假設 $a_x=1/2$ (U.D.D.)

其中 $$L_x = \int_0^1 l_{x+t}\, dt = l_{x+1} + a_x d_x$$

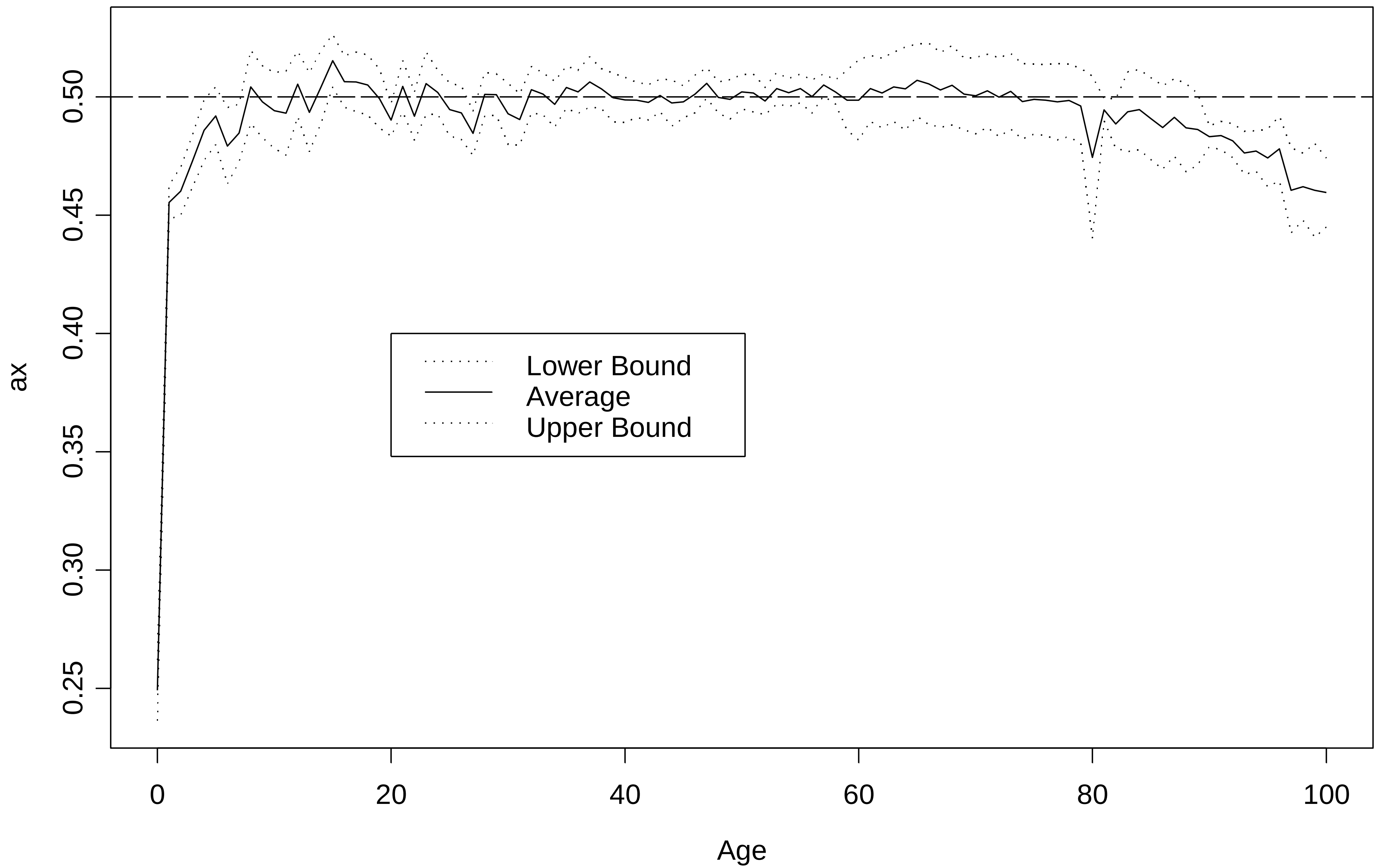
→ 根據 Chiang(蔣慶琅,1984) 對世界26個國家的計算結果，在0至4歲都發現 $a_x < 1/2$，尤其是0歲時最明顯。

不同死亡假設下的終壽區間成數a_x

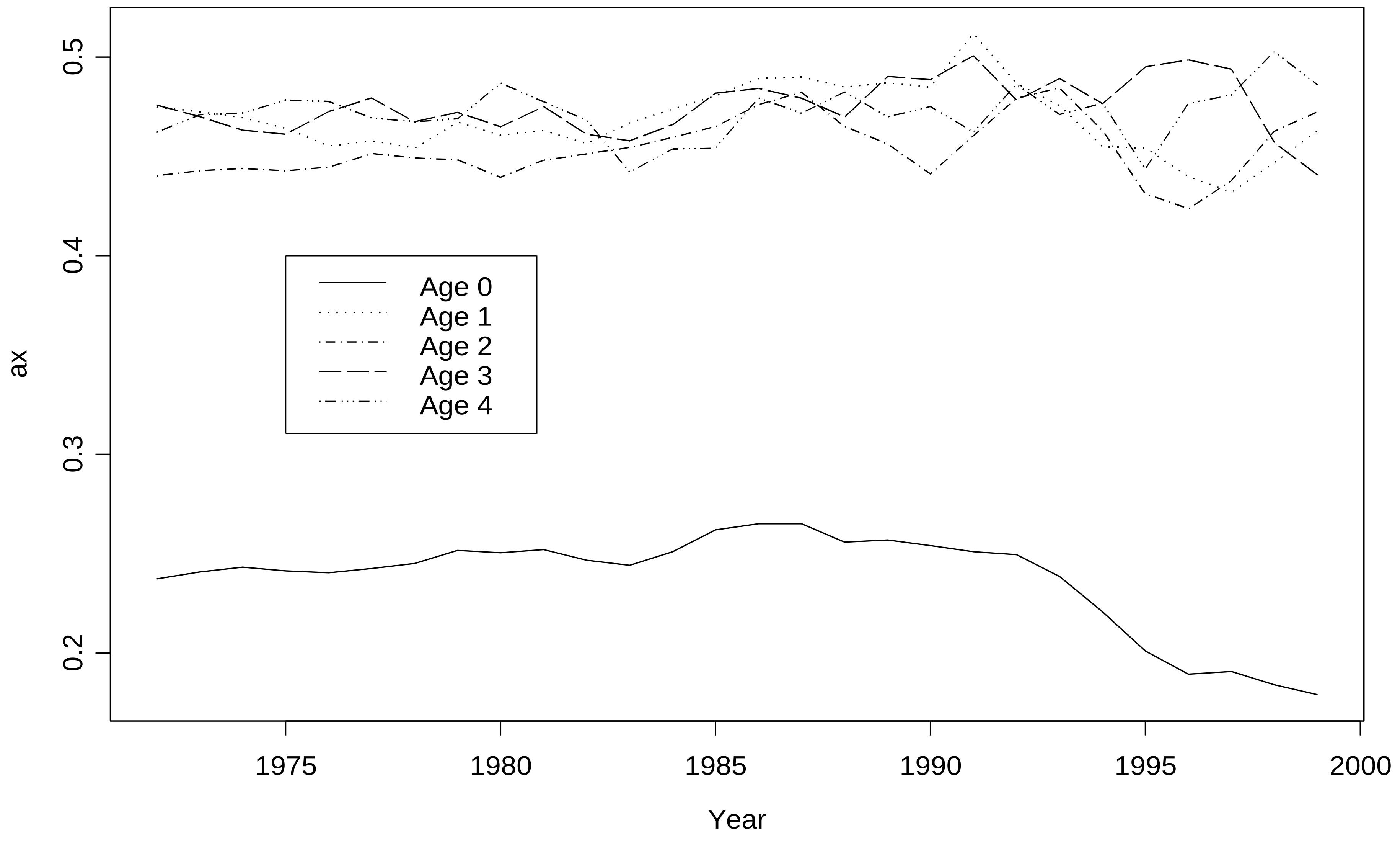




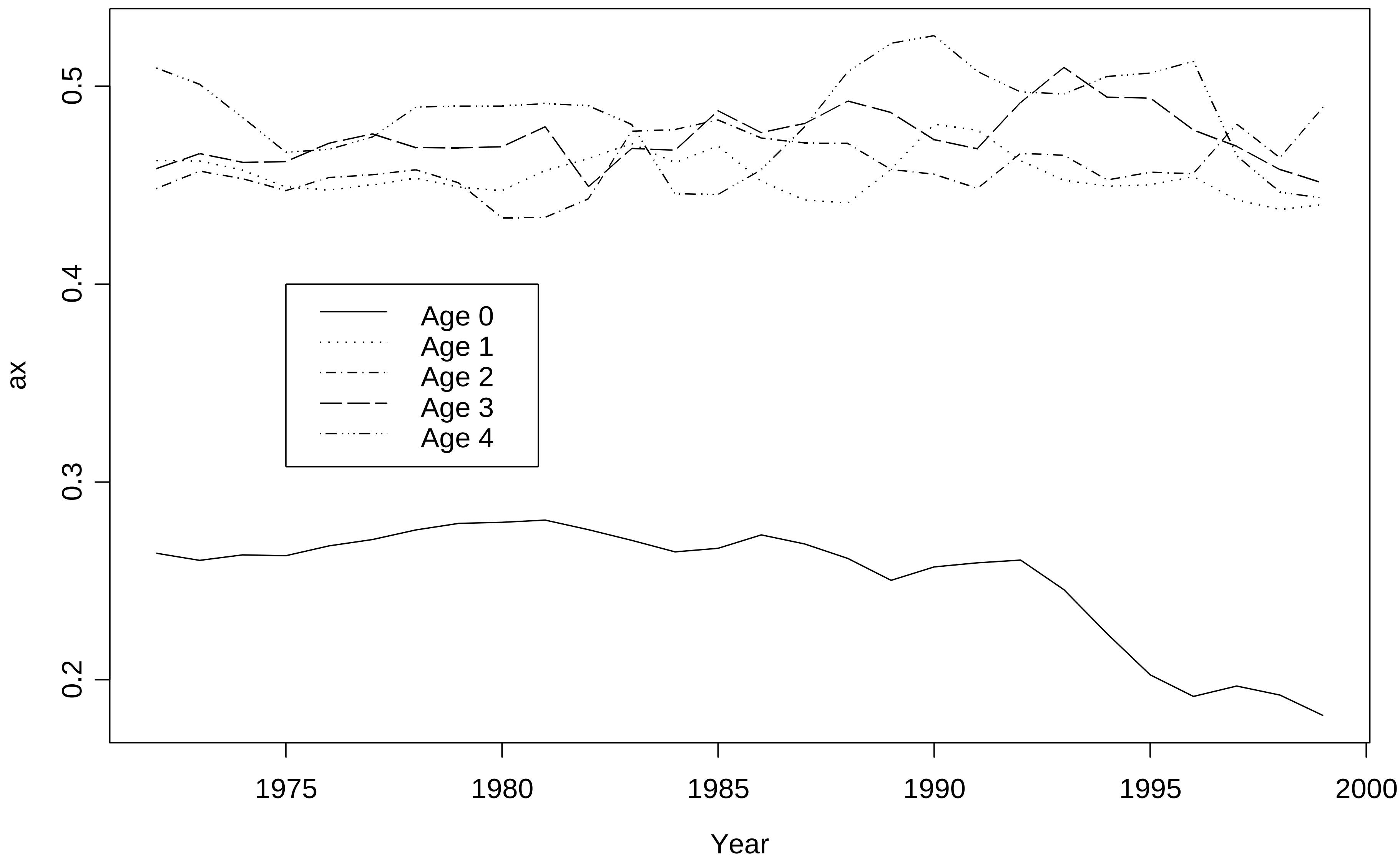
臺灣1971至2000年各年齡終壽區間成數(男性)



臺灣1971至2000年各年齡終壽區間成數(女性)



1971～2000年0-4歲終壽區間成數(男性；3年移動平均)



1971～2000年0-4歲終壽區間成數(女性；3年移動平均)

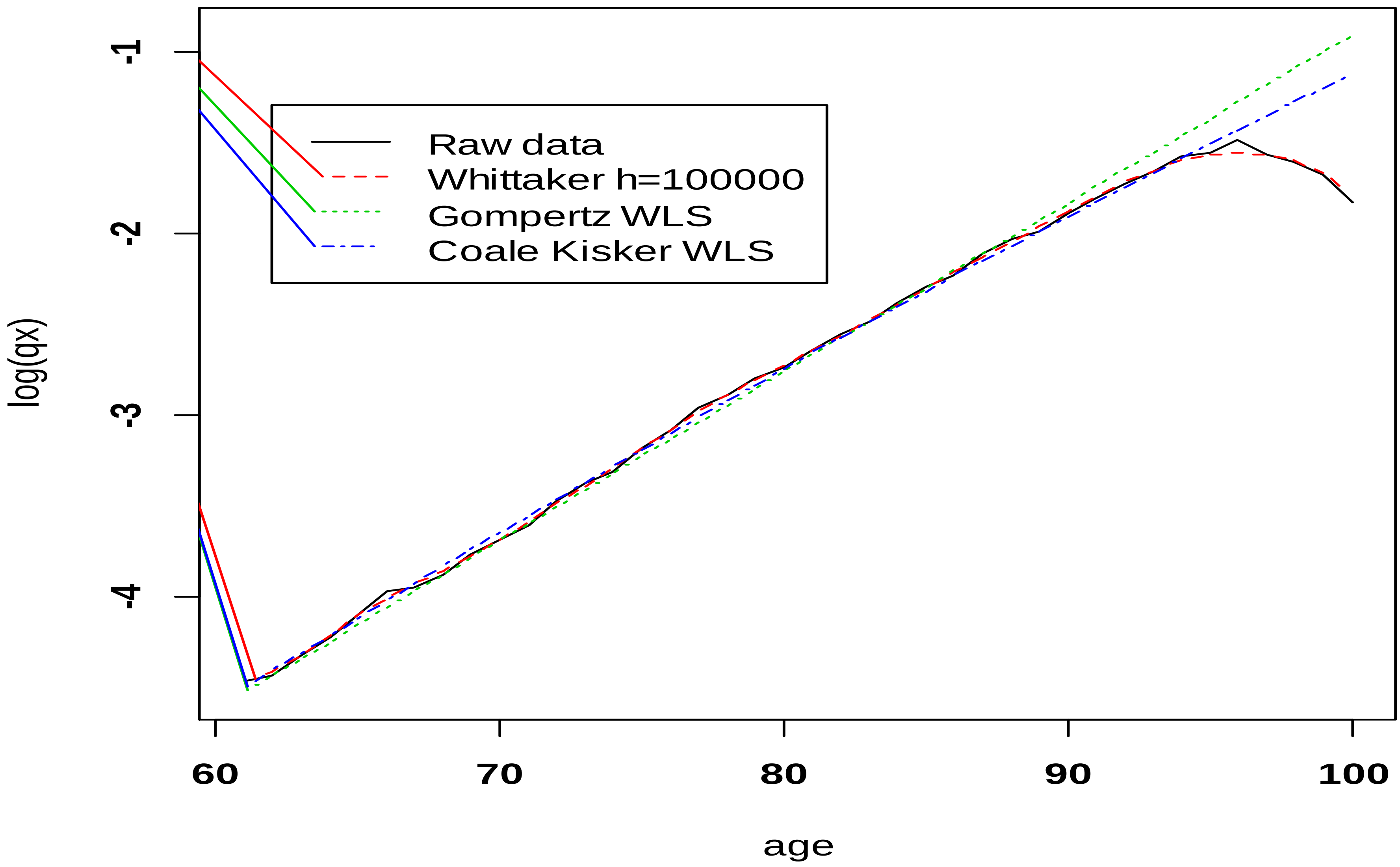
- 1歲至4歲的 a_x 在民國60年至89年間大致維持定值。
- 0歲的終壽區間成數則呈現不同的趨勢，在民國80年前的走勢大致穩定，民國80年後則逐漸下降。

→兩者有顯著差異；

→民國60至79年的 a_0 約為0.25；

→民國80至89年的 a_0 約為0.21。

(但與WHO建議的數值仍有差異。)



2009-2011年臺灣高齡男性死亡率修勻結果

關於修勻方法的想法

- 如果人數（或是暴露數）足夠，修勻可能改變資料的原有特性。

→年輕年齡組、高齡族群因為人數或死亡人數較少，需要資料調整的方法。

- 傳統的修勻大多合併相鄰年齡的人口資料，以「增加樣本數」的角度，降低死亡率及發生率的震盪幅度。

註：這類型的方法可稱為「同地同時」。

關於合併資料

→<u>**同地同時**</u>：一般修勻方法（合併年齡）。

→<u>**同地異時**</u>：合併三年或多年資料，或套用死亡率模型，像是Lee and Carter (1992)，或是以歷史資料結合修勻方法。

→<u>**異地同時**</u>：也可參考全國死亡率，例如：Whittaker比值法、Partial SMR(標準死亡比)、貝氏修勻等方法。

→<u>**異地異時**</u>：代入不同各縣市資料，會因地區特性產生參數估計偏誤，可用Lee-Carter連貫模型(Li and Lee, 2005)及加入修勻方法修正。

「同地同時」的修勻方法

- 「同地同時」使用編算地區一年度的資料，透過合併相鄰年齡增加人數，方法包括：

→Moving Weighted Average (MWA)

→Whittaker-Henderson

→Kernel（核）

→Spline（花鍵、樣條）

→Curve Fitting (e.g., Local Poly, GAM, Wavelet)

→Parametric Models (e.g., Gompertz, Logistic, Coale-Kisker, Heligan-Pollard, Wang-Yue)

「同地異時」、「異地同時」

- 「同地異時」與「異地同時」皆參考其他資訊，像是該地區的歷史趨勢，或是類似地區的同一年資料。

→Bayesian Analysis

→Whittaker Ratio/Partial SMR

→Lee-Carter Model (or PCA)

→Reduction Factor（英國CMI）

→APC & other parametric models (e.g. CBD)

「異地異時」的修勻方法

- 「異地異時」同時使用不同編算地區、不同年度的資料，需考量資料的同質性，這類型的方法多半以參數模型為主。

→Bayesian Lee-Carter Model

→Li-Lee Model (Coherent Lee-Carter) Model

→Lee-Carter Model + Partial SMR (or Whittaker Ratio)