

# 統計計算與模擬

政治大學統計系余清祥

2025年3月18日~3月25日

第四單元：模擬檢定

<http://csyue.nccu.edu.tw>





# Permutation tests and Bootstrap

- 模擬檢定的部份將介紹

- Permutation Tests

- Monte Carlo Tests (and p-values)

- Bootstrap and Resampling Methods

- Preliminary Exploration of Unknowns

- Applications



# 方法上的差異

## ■ Permutation Tests

→ 找出資料的所有可能排列或組合

## ■ Monte Carlo Tests (Monte Carlo p-value)

→ 在模型假設下產生亂數

## ■ Bootstrap and Resampling Methods

→ 對資料執行隨機抽樣(Nonparametric)

→ 以模型假設配合資料求出參數，再產生亂數(Parametric)

# Monte Carlo vs. Bootstrap

■ 蒙地卡羅(Monte Carlo)與拔靴法(Bootstrap)兩者操作方式類似：

- 蒙地卡羅認為母體分配已知，從這個假設產生亂數，再從這些亂數探討問題，包括期望值、變異數、百分位數等母體性質；
- 拔靴法「假設」樣本就是母體，從這個有限母體中產生亂數，再從這些亂數計算樣本統計量的「變異數」！

註：拔靴法的功能較為侷限。

# 模擬應用範例：

## ■ Monte Carlo Sampling:

→ 檢查抽出的樣本是否隨機。

應用範例：

→ 是否應該聽從某股市名嘴的建議，投資某些「熱門股票」？下表中的平均投資報酬為7.05% [=(\$468.873-\$438)/\$438]。

→ 問題：這個報酬是否明顯高於任意選股？

**Table 3.1**  
**The stocks picked by the analyst**

	<u>Price at</u> <u>12/31/85</u>	<u>Price at</u> <u>12/31/86</u>	<u>Dividend</u> <u>for 1986</u>
PPG INDUSTRIES INC	51.000	72.875	1.880
HILTON HOTELS CORP	64.875	67.250	1.800
GENERAL ELECTRIC CO	72.750	86.000	2.370
MCDERMOTT INTL INC	18.250	21.750	1.800
LOUISIANA LAND & EXPLORATION	30.250	27.250	1.000
DRESSER INDUSTRIES INC	18.125	19.375	0.700
MCDONALD'S CORP	80.875	60.875	0.645
WEST POINT-PEPPERELL	43.375	52.125	2.385
SERVICE CORP INTERNATIONAL	31.250	24.750	0.293
AMERADA HESS CORP	<u>27.250</u>	<u>23.750</u>	<u>0.000</u>
Total	438.00	456.00	12.873

隨機抽取10支股票，在1000次的模擬中，  
有26%(260次)的機會報酬率不小於7.05%。  
→ 與隨機選股無異！！

■ 另一個範例：(關於刪除離群值的迷思！)

→ 資料分析若發現離群值，直接刪去不見得是最好的辦法，因為離群值可能蘊含珍貴的資訊。

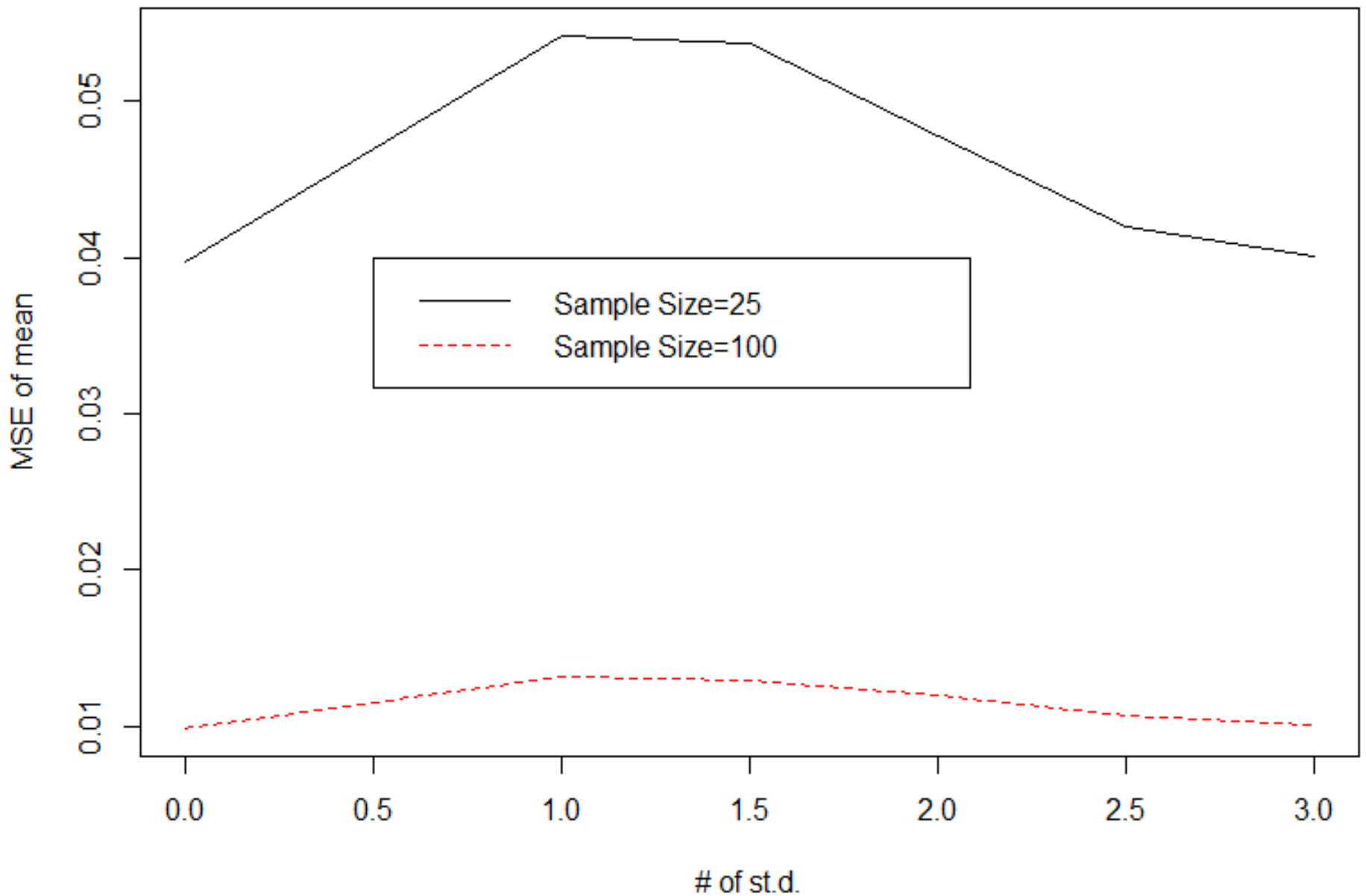
→ 假設觀察值服從標準常態分配，我們的目標在估計期望值及變異數。比較以下兩種方法的精確度(Mean Squared Error)：

一、樣本平均數（記為 $a=0$ ）；

二、去掉 $a$ 倍標準差後的平均數（1, 1.5, 2, 2.5, 3總共五種情境）。

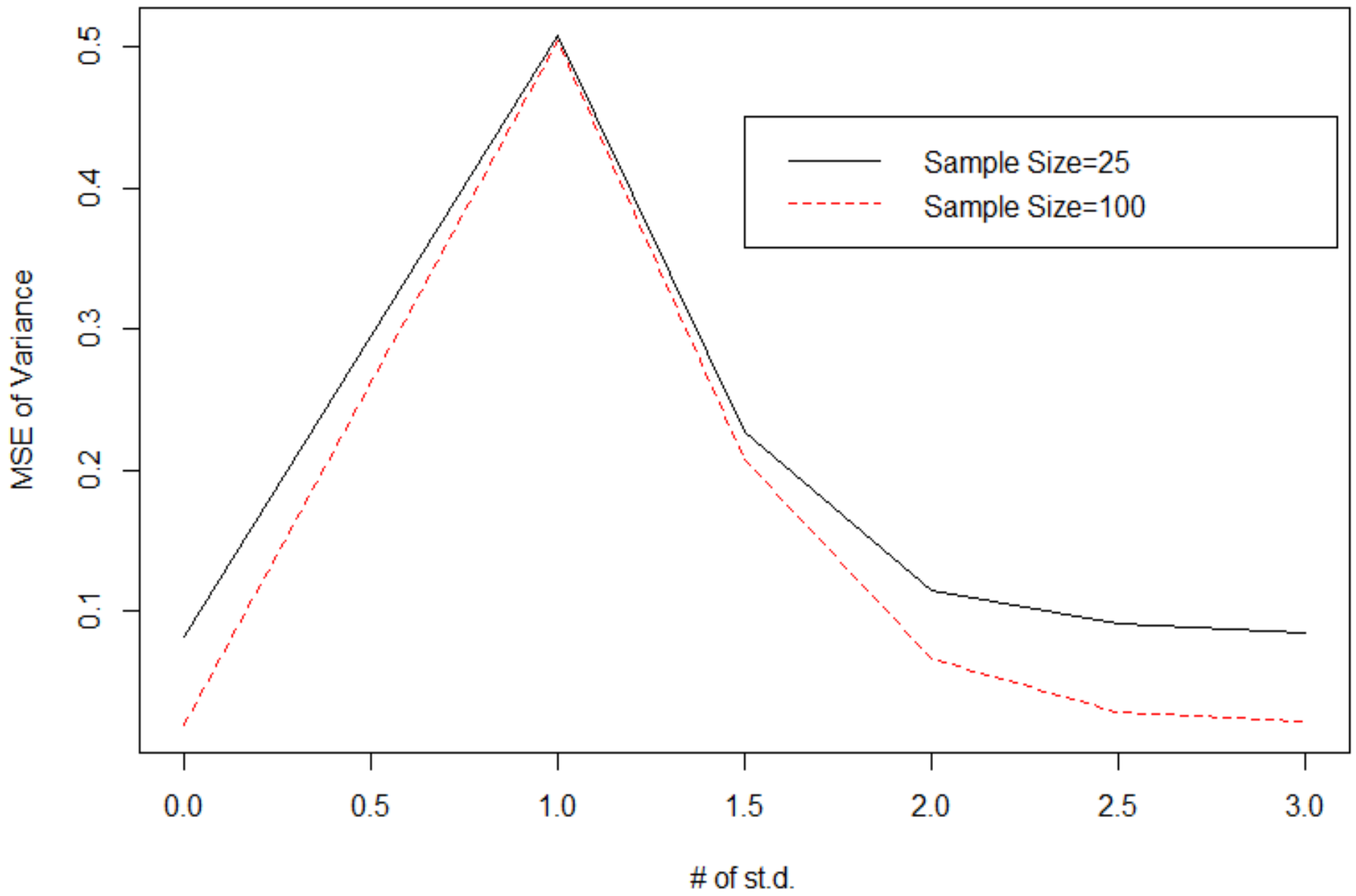
假設觀察值有25及100個，模擬10,000次。

# 不同期望值估計值的MSE比較





# 不同變異數估計值的MSE比較

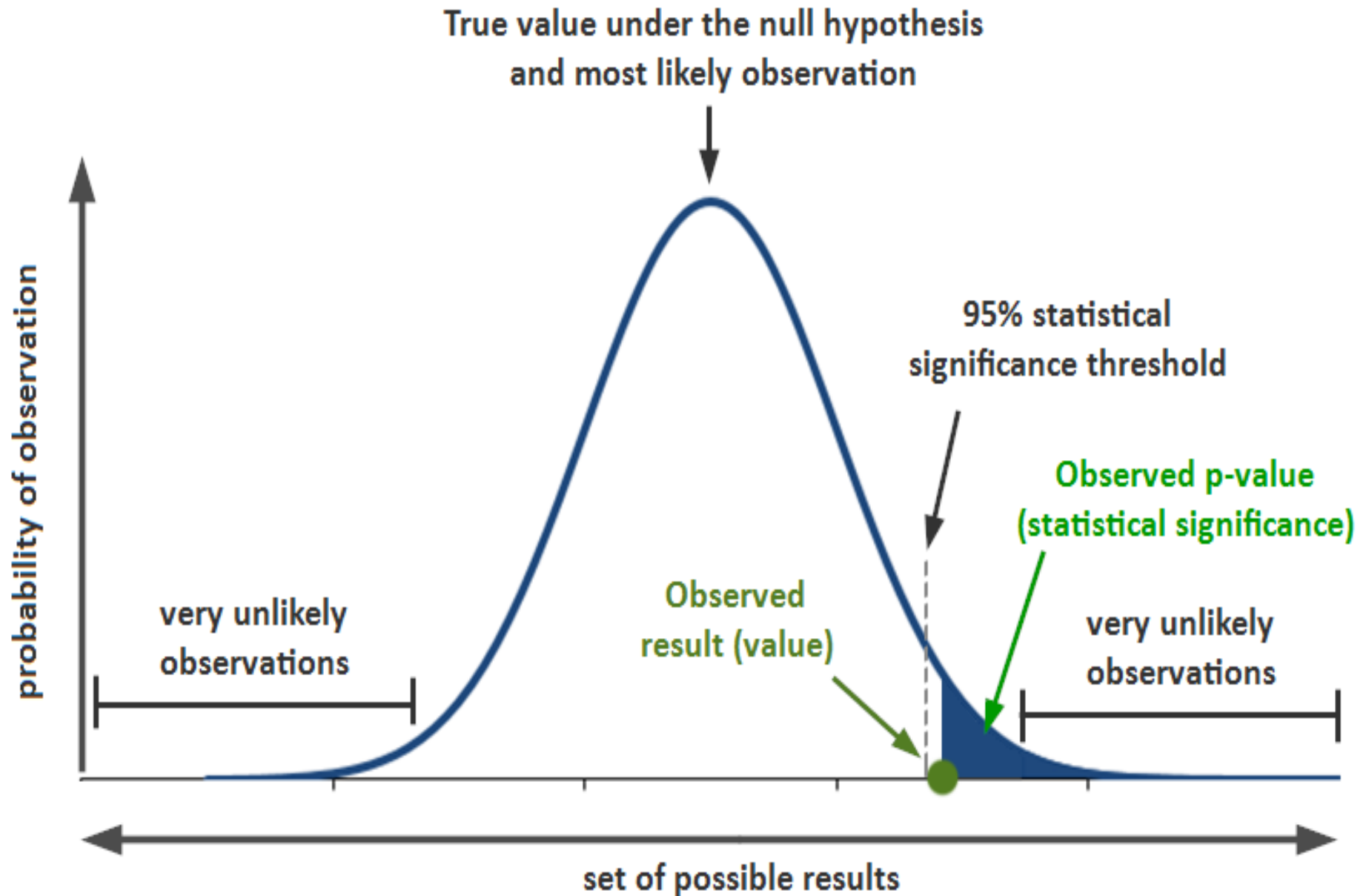


```

t1=NULL
t2=NULL
for (i in 1:10000) {
  x=rnorm(100)
  a=mean(x)
  b=sqrt(var(x))
  x1=x[x > a-b & x < a+b]
  x1.5=x[x > a-1.5*b & x < a+1.5*b]
  x2=x[x > a-2*b & x < a+2*b]
  x2.5=x[x > a-2.5*b & x < a+2.5*b]
  x3=x[x > a-3*b & x < a+3*b]
  y1=c(mean(x),mean(x1),mean(x1.5),mean(x2),mean(x2.5),mean(x3))
  y2=c(var(x),var(x1),var(x1.5),var(x2),var(x2.5),var(x3))
  t1=cbind(t1,y1)
  t2=cbind(t2,y2)
}
matplot(xx,yy,type="l",lty=c(1,2),col=c(1,2),xlab="# of st.d.",ylab="MSE of
mean")
legend(0.5,0.04,c("Sample Size=25","Sample Size=100"),lty=c(1,2),col=c(1,2))
matplot(xx,yz,type="l",lty=c(1,2),col=c(1,2),xlab="# of st.d.",ylab="MSE of
Variance")
legend(1.5,0.45,c("Sample Size=25","Sample Size=100"),lty=c(1,2),col=c(1,2))

```

# 統計檢定與p-value



## ■ 統計檢定與p-value

→ p-value測量觀察值的「罕見程度」，數值愈小、出現機會愈小。

■ 根據「虛無假設」及「檢定量」描繪出可能狀況，如：t-test之類的檢定。

→ 問題：如果缺乏檢定量的資訊，或是不確定假設條件是否成立，如何檢測觀察值？

→ 排列檢定（均勻分配）、t檢定（常態分配）是兩種常見的假設。

註：「顯著性」也與假設條件有關。

# Concept of Permutation Test

## Structure



## Random Pattern



## ■ Permutation test:

→ 以排列組合的方式計算事件的發生機率，常見於兩組樣本的比較，尤其當資料不是常態分配時，或是樣本數較少時。(使用R的BHH2模組之「permtest」。)

### ■ 檢定步驟：

- 定義問題、確認虛無與對立假設；
- 確定檢定量；
- 由觀察值計算檢定量；
- 重複抽樣獲得所有可能的檢定量；
- 由排列檢定的結果決定檢定結果。



應用範例：

→ 某汽油添加劑宣稱可增加每公升的里程數，以下為各測試4次添加與不添加的實驗結果。

◆ 添加：33.0, 31.0, 34.5, 34.0

◆ 不添加：29.5, 32.0, 32.9, 31.5

→ 因為樣本數少，除非差異較大，檢定結果通常是不能拒絕兩者的期望值相等。  
(t-檢定的 p-value 等於0.170。)

→8個觀察值任意抽出4個，所有可能為：

$$\binom{8}{4} = \frac{8 \times 7 \times 6 \times 5}{4 \times 3 \times 2 \times 1} = 70 \text{ 種}$$

若藉由Permutation test，知道有6種可能使得兩者的差異不小於觀察差異(即差異為6.6)。因此， $p\text{-value} = 6/70 = 0.086$ 。

註：排列檢定屬於無母數方法，在觀察值個數不多時的其檢力通常較小。





→ 以R指令「permtest」測試

```
x1=c(33.0, 31.0, 34.5, 34.0)
```

```
y1=c(29.5, 32.0, 32.9, 31.5)
```

```
permtest(x1,y1)
```

<i>N</i>	<i>t.obs</i>	<i>t-Dist:P(&gt;t)</i>	<i>PermDist:P(&gt;t)</i>
70.00000000	1.56172013	0.08468935	0.07142857

→ 藉由手算Permutation test，知道有6種可能使得兩者的差異不小於觀察差異(即差異為6.6)。因此， $p\text{-value} = 6/70 = 0.086$ 。

■ 再一個範例：

比較兩組樣本的變異程度，Good(1994)建議轉換原始樣本為  $X_i' = |X_i - \text{median}(X)|$ ，

再計算  $S = \sum_{i=1}^{n-1} X_i'$ ，其中  $X_n = \text{median}(X)$ 。

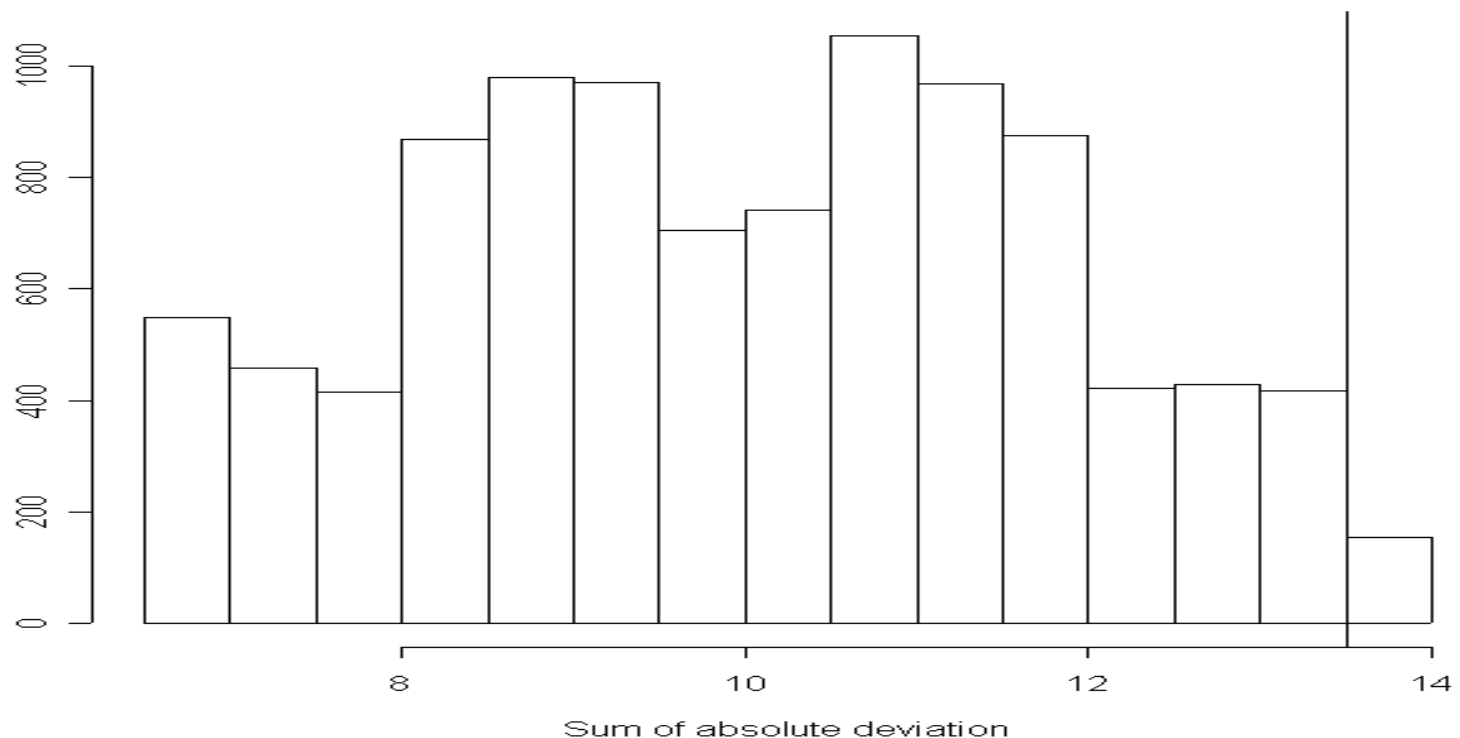
原則上， $S$  愈大變異也就愈大。

→ 問題：何種程度的  $S$  的差異才算有差別？

範例：兩組樣本  $X = \{121, 123, 126, 128.5, 129\}$   
及  $Y = \{153, 154, 155, 156, 158\}$ 。

→ 計算得出  $X' = \{5, 3, 2.5, 3\}$ 、 $S_X = 13.5$ 、  
 $Y' = \{2, 1, 1, 3\}$ 、 $S_Y = 7$ 。

- 以  $X'$  及  $Y'$  的共 8 個數值為依據 (8 中抽 4)，看看模擬的數值中有幾次的  $S$  值不小於 13.5。
- 模擬 100,000 次，發現  $S$  值大於或等於 13.5 的模擬次數有 5712 次， $p$ -value 等於 0.05712，表示  $X$  的變異程度大於  $Y$ 。(排列檢定 = 4/70)

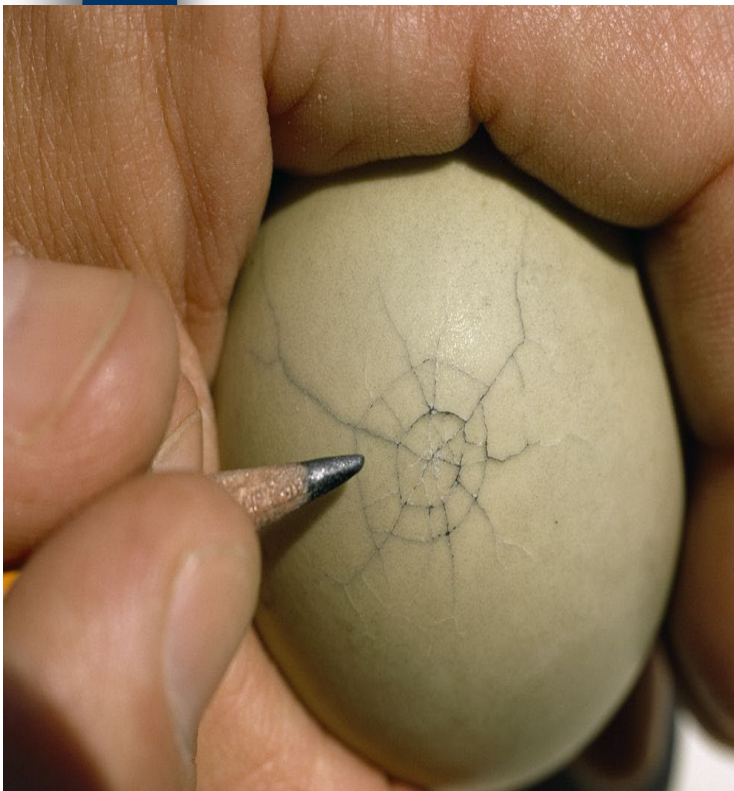


- 另一個範例：(兩個變數間是否相關！)如  
如果只有兩個變數  $(X, Y)$ ，但觀察值的個  
數不多、或是變異過大，藉由相關係數  
得出的結果一般都不顯著。

→ 計算所有可能的  $S = \sum_{i=1}^n x_i y_{\phi(i)}$  數值，看看  
有幾次比觀察值大(或小)，其中  $\phi$  為定義  
於  $\{1, 2, \dots, n\}$  的排列。

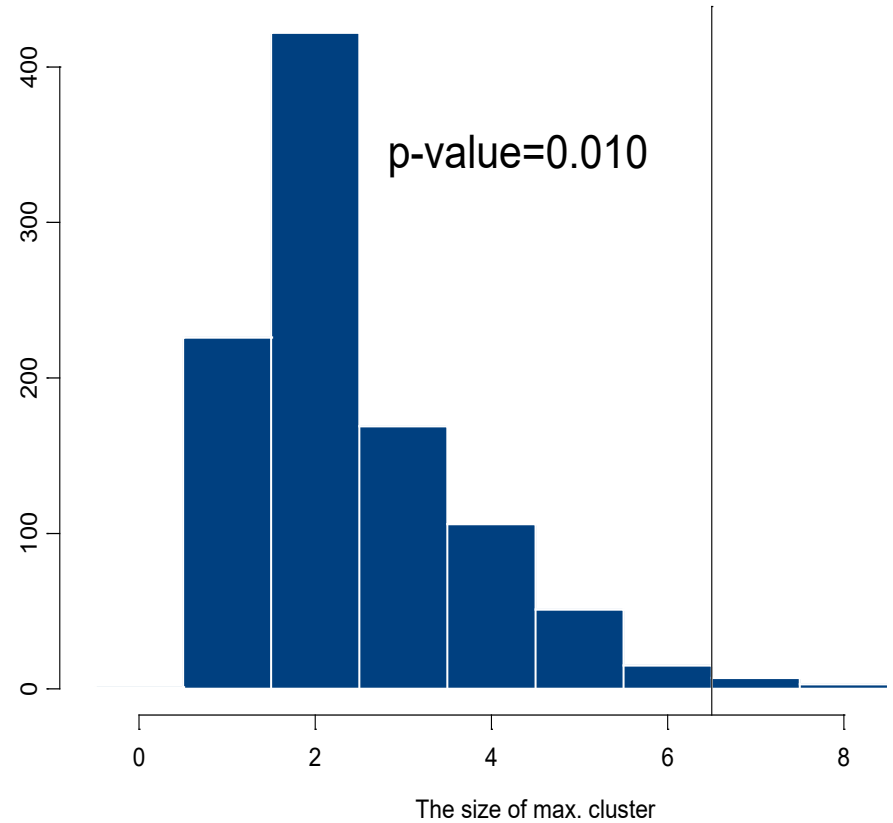
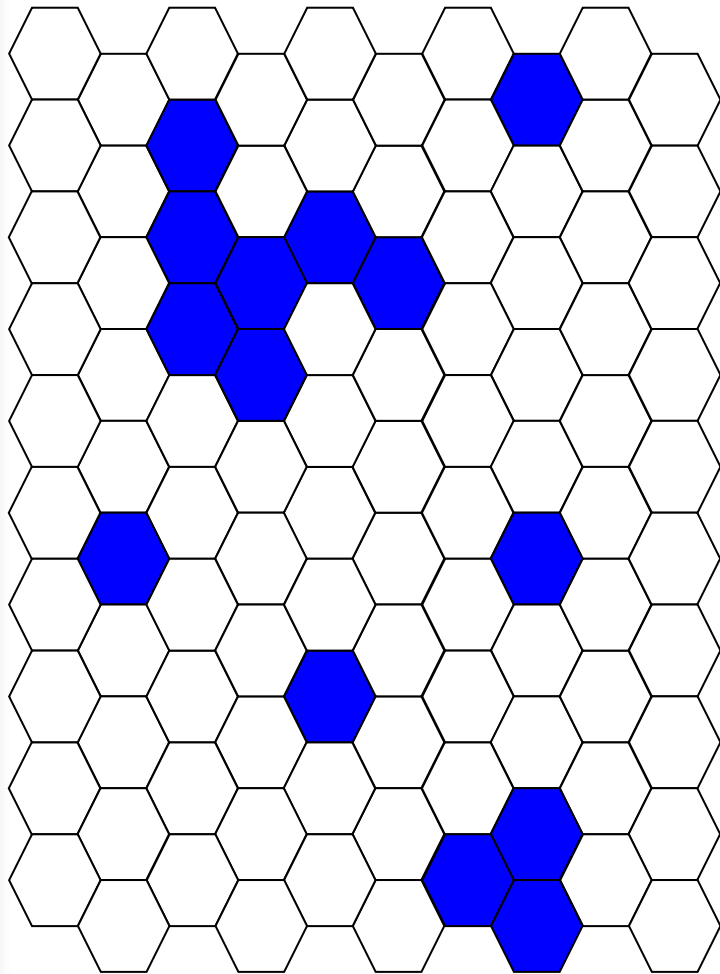


	1	2	3	4	5
DDT (ppm)	65	98	117	122	130
Egg Shell Thickness	.52	.53	.49	.49	.37



- 如果直接代入計算，在常態分配的假設下得出相關係數  $-0.663$ ， $p$  值  $0.223$ 。
- 若以排列檢定，因為 120 種組合中僅有兩種的  $S$  值不大於觀察值的  $S$  值，表示兩個變數間存在非常強的負相關。

# 檢查是否有群聚(Cluster)的現象！



An example of testing the largest cluster (1,000 runs)

# 霧霾出現的天數是否為連續(群聚)?

日	一	二	三	四	五	六
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

註：31天中有10天有霧霾，連續四天算是群聚嗎？(Run test?)

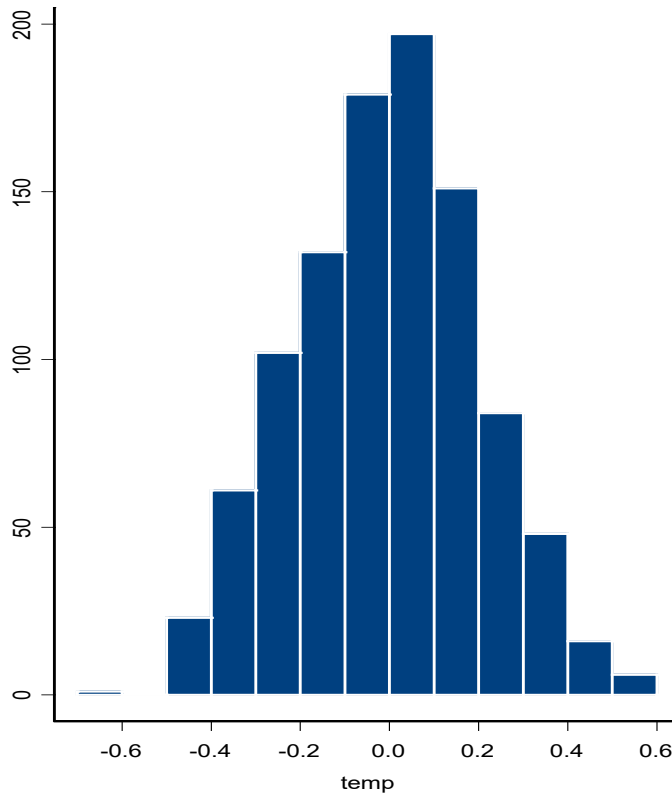
# Monte Carlo p-value & Test

- 根據虛無假設模擬出  $n$  組樣本，如果觀察值排在第  $k$  個(也就是第  $k$  大)，則檢定的  $p\text{-value} = k/(n+1)$ 。一般顯著水準為 0.05， $n$  值多半選為 99、499、999、9,999 或 99,999 等數字。
- Q：蒙地卡羅 p-value 與排列檢定的差別？
  - 蒙地卡羅檢定的作法類似，給定模型假設下，求出統計量等相關數值的分配。
  - 應用範例：迴歸分析中有人用調整過的  $R^2$ ，以消除因隨機而造成的線性相關。

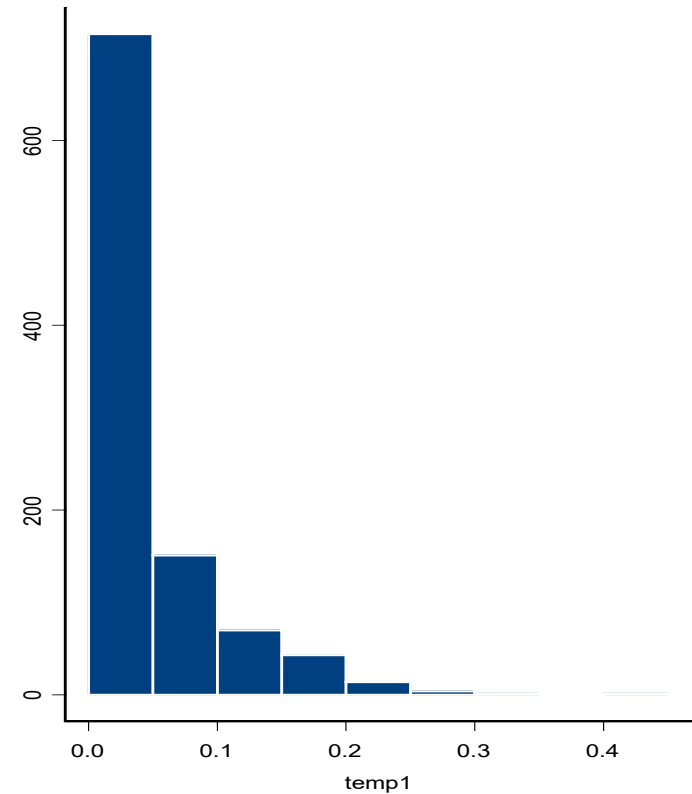


- 隨機由  $N(0,1)$  產生互相獨立的25個  $X$ 、 $Y$  觀察值，根據迴歸方程式  $Y = \alpha + \beta X$  計算  $R^2$ ，重複1000次的模擬可得

Correlation between X and Y



R<sup>2</sup> of Y on X



→ 平均  $R^2 = 4.2\%$ ， $Q3 = 5.9\%$ 。



Note: The relationship between  $R^2$  and  $R_a^2$  is

$$R_a^2 = 1 - (1 - R^2) \left( \frac{n-1}{n-p} \right).$$

Or, equivalently,

$$1 - R_a^2 = (1 - R^2) \left( \frac{n-1}{n-p} \right).$$

Plugging into  $n = 25$  and  $p = 2$ , we get

$$R_a^2 = 1 - (1 - 4.2\%) \times \frac{24}{23} = 0.035\% \cong 0.$$



■ 另一個範例：

迴歸分析中常以殘差的符號變化，判斷殘差間是否互相獨立。例如：如果共有20個殘差，其正負號為

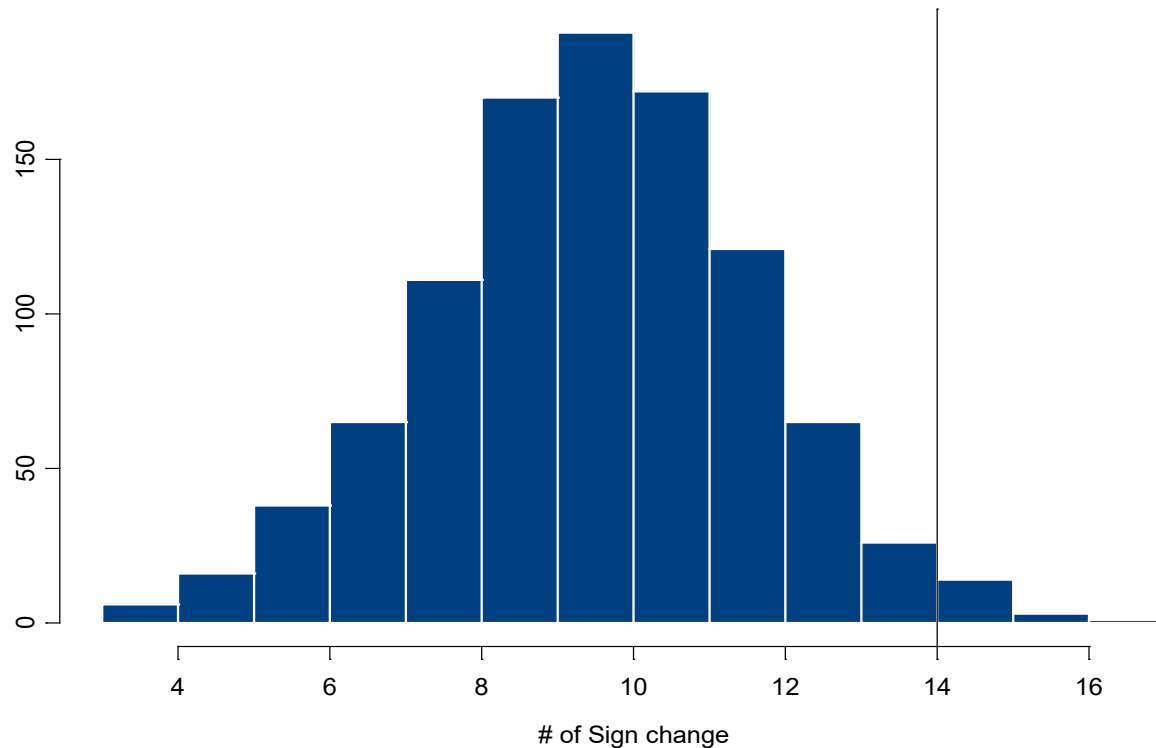
+ - - + - + - + + - + - + + - - + - - +

共有14個符號改變，想檢查是否殘差間是否為負相關。

→ 符號改變頻繁，表示負相關；

→ 符號不改變，表示正相關。

→ 以10個正號、10個負號為基礎，重複模擬999次，計算符號改變的次數。發現大於或等於14次符號改變的模擬次數有44次，因此  $p\text{-value} = (44+1)/(999+1) = 0.045$ ，也就是殘差間呈現負相關。



■ 另一個範例： $X_i \sim i.i.d. N(m, \sigma^2), i = 1, 2, \dots, n.$

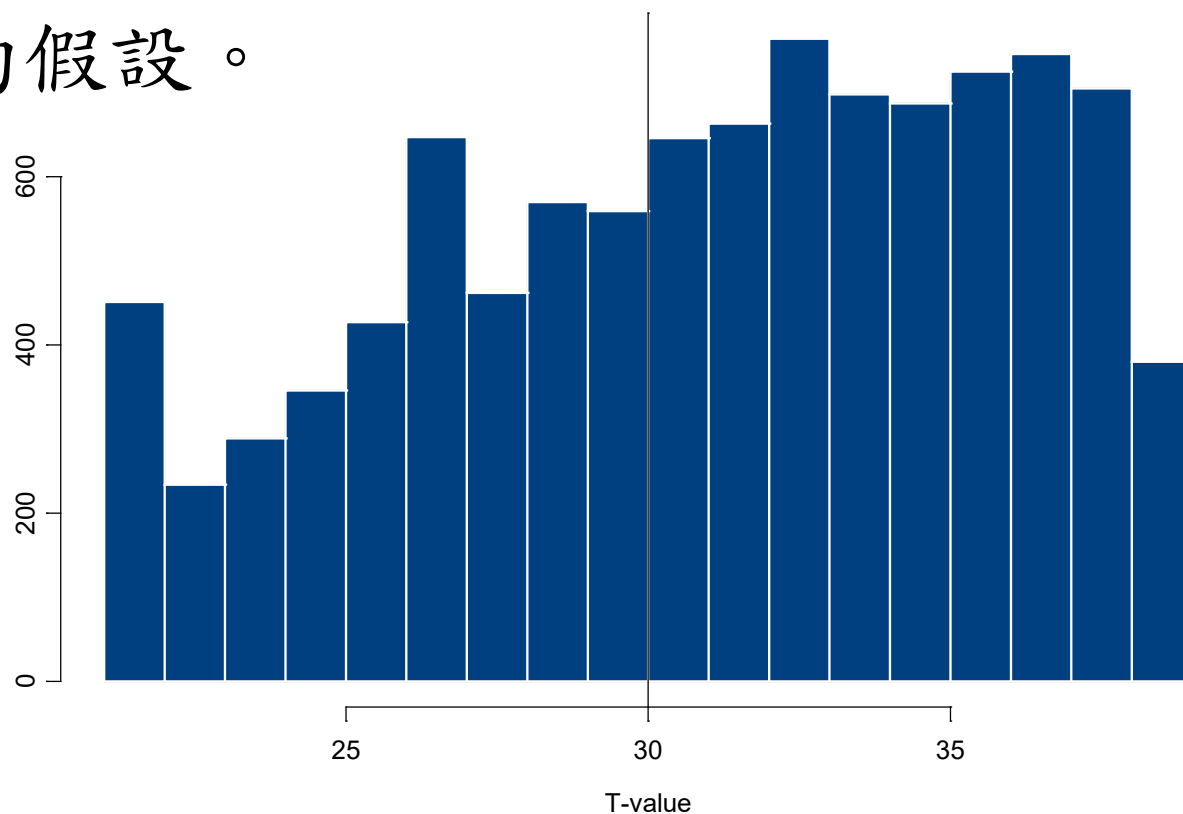
Wilcoxon符號檢定可用來檢查一組樣本的中位數是否等於某個定值，首先計算所有觀察值的秩(rank)，再乘以大於或小於中位數的符號，分別加總正負兩個秩( $T^+$ 、 $T^-$ )的總和，取較小者為統計檢定量。

→ 理論上， $T^+ + T^- = \frac{n(n+1)}{2}$ 。

→ 當樣本數較大時，

$$E(T) = \frac{n(n+1)}{4} \quad \& \quad Var(T) = \frac{n(n+1)(2n+1)}{24}.$$

→ 以  $X_i \sim i.i.d. N(0,1), i = 1, 2, \dots, 12$  為例，假設某組觀察值的檢定量為30，重複9,999次的模擬，計算出Monte-Carlo p-value等於  $(1+3985)/(1+9999)=0.3986$ ，不拒絕中位數為0的假設。



Critical values of the Mann-Whitney-Wilcoxon test ( $\alpha=.05$ , 10,000 runs, and numbers in red are the true critical values.)

|         | $n_1=2$ | 3      | 4      | 5  | 6  | 7  | 8      | 9      | 10 |
|---------|---------|--------|--------|----|----|----|--------|--------|----|
| $n_2=2$ | 3       | 3      | 3      | 3  | 3  | 3  | 4      | 4      | 4  |
| 3       | 6       | 6      | 6      | 7  | 8  | 8  | 9      | 9      | 10 |
| 4       | 10      | 10     | 11     | 12 | 13 | 14 | 14(15) | 16(15) | 16 |
| 5       | 15      | 16     | 17     | 18 | 19 | 21 | 22     | 23     | 24 |
| 6       | 21      | 23     | 24     | 25 | 27 | 28 | 30     | 31(32) | 33 |
| 7       | 28      | 30     | 32     | 34 | 35 | 37 | 39     | 41     | 43 |
| 8       | 37      | 39     | 41     | 43 | 45 | 47 | 49(50) | 52     | 54 |
| 9       | 46      | 48     | 51(50) | 53 | 56 | 58 | 61     | 63     | 66 |
| 10      | 56      | 58(59) | 61     | 64 | 67 | 70 | 73     | 76     | 79 |

# Bootstrap(拔靴法)

- Bootstrap法可追塑至Efron於1979提出的方法，屬於重複抽樣(Resampling)方法。將已有的觀察值當作是母體重複抽樣(與Monte Carlo有真實母體不同！！)，以求取原先因資料不足而無法探討的資料特性，早期探討的特性以變異數為主。
- 舉例而言，假設  $x_1, x_2, \dots, x_N$  為來自同一分配的觀察值，而我們想瞭解這個分配的中位數與其中位數的變異數。



範例：一組由Poisson分配抽出的隨機樣本

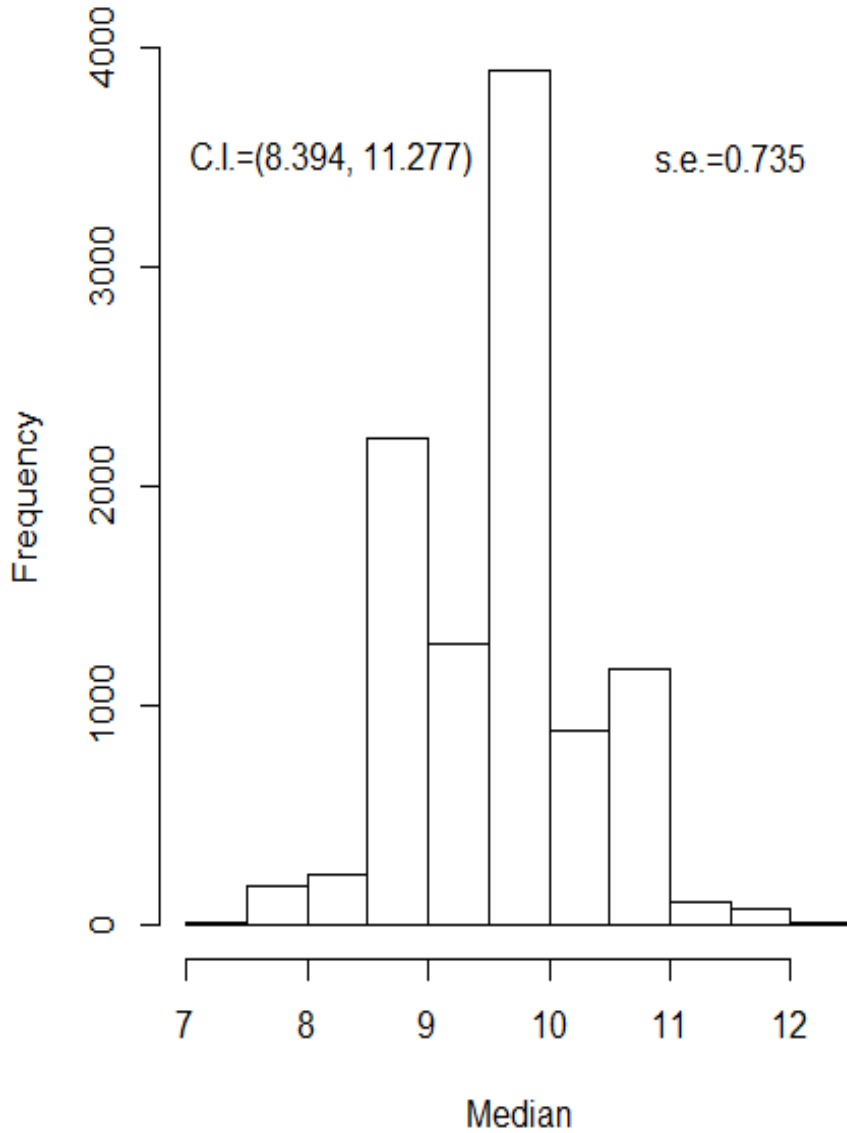
6 7 7 7 7 8 8 8 9 9 9 9 10 10 10 10  
11 11 12 13 13 13 13 14 15 15 15 15 17 20

→ 已知樣本中位數為10，欲求出母體中位數的信賴區間，可由Bootstrap模擬出標準差。10,000次模擬得出標準差估計值1.0063，中位數的95%信賴區間 (8.5586, 11.4414)

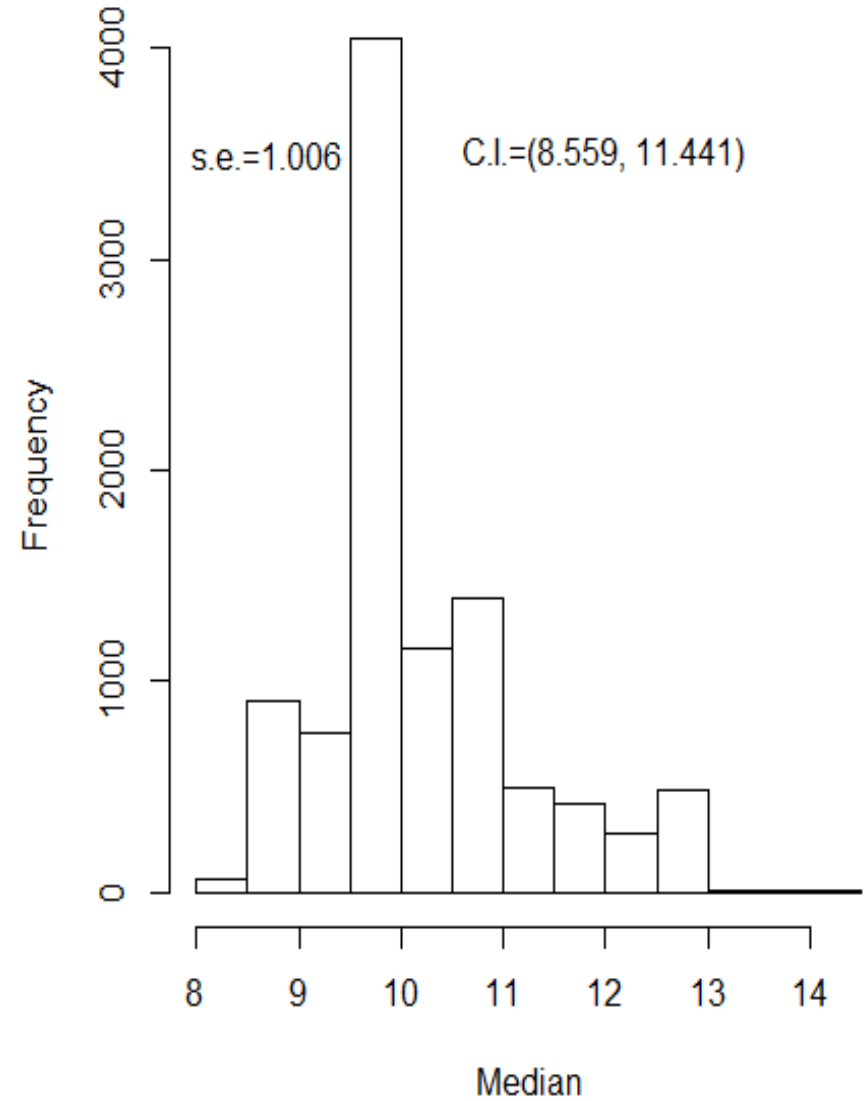
涵蓋了實際中位數10。

註：母體為Poisson(10)；蒙地卡羅95%信賴區間為(8.3939, 11.2767)。

Monte Carlo (10,000 runs)



Bootstrap (10,000 runs)

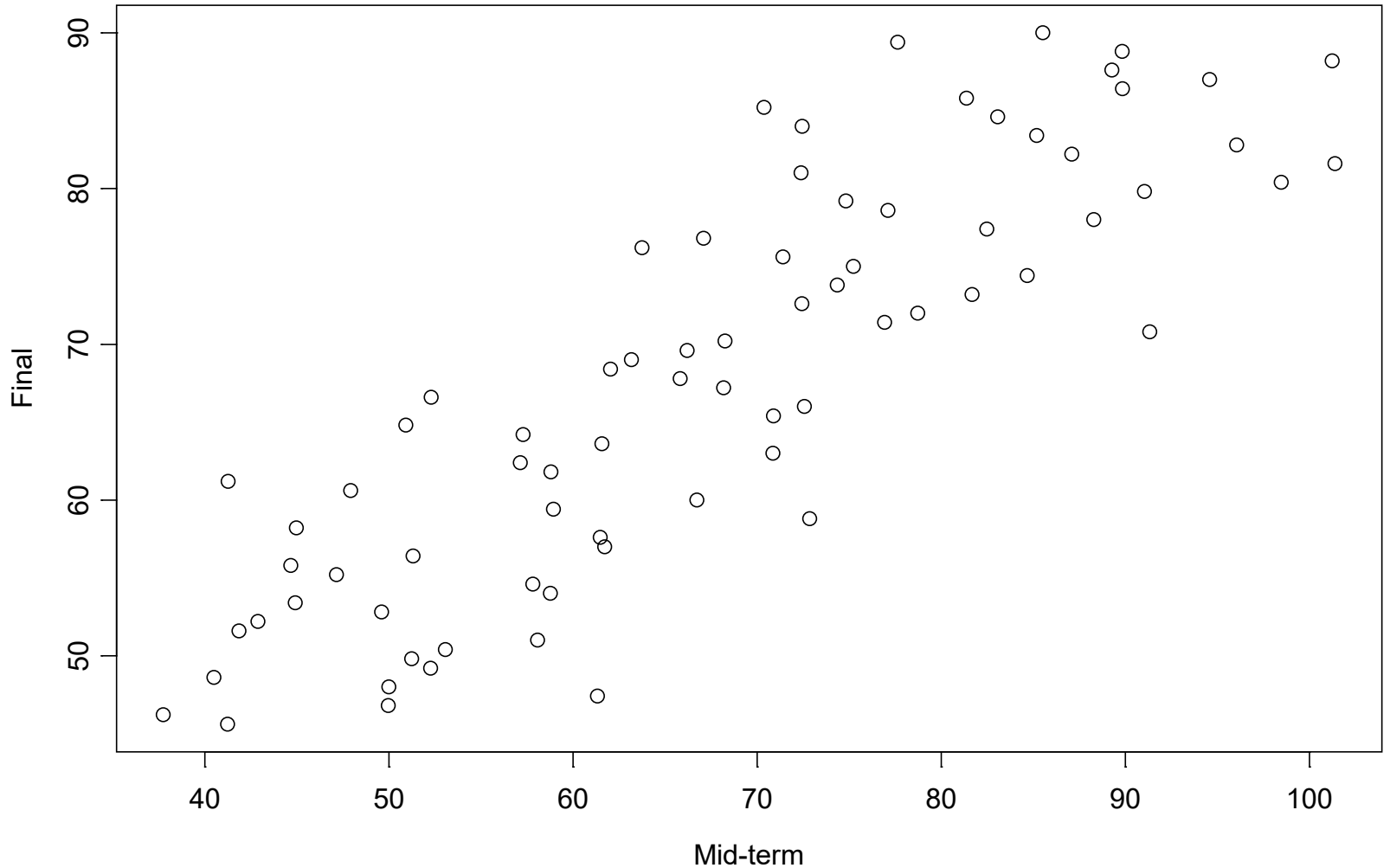


兩者的差異並不大！(標準誤0.7354 vs. 1.0063)

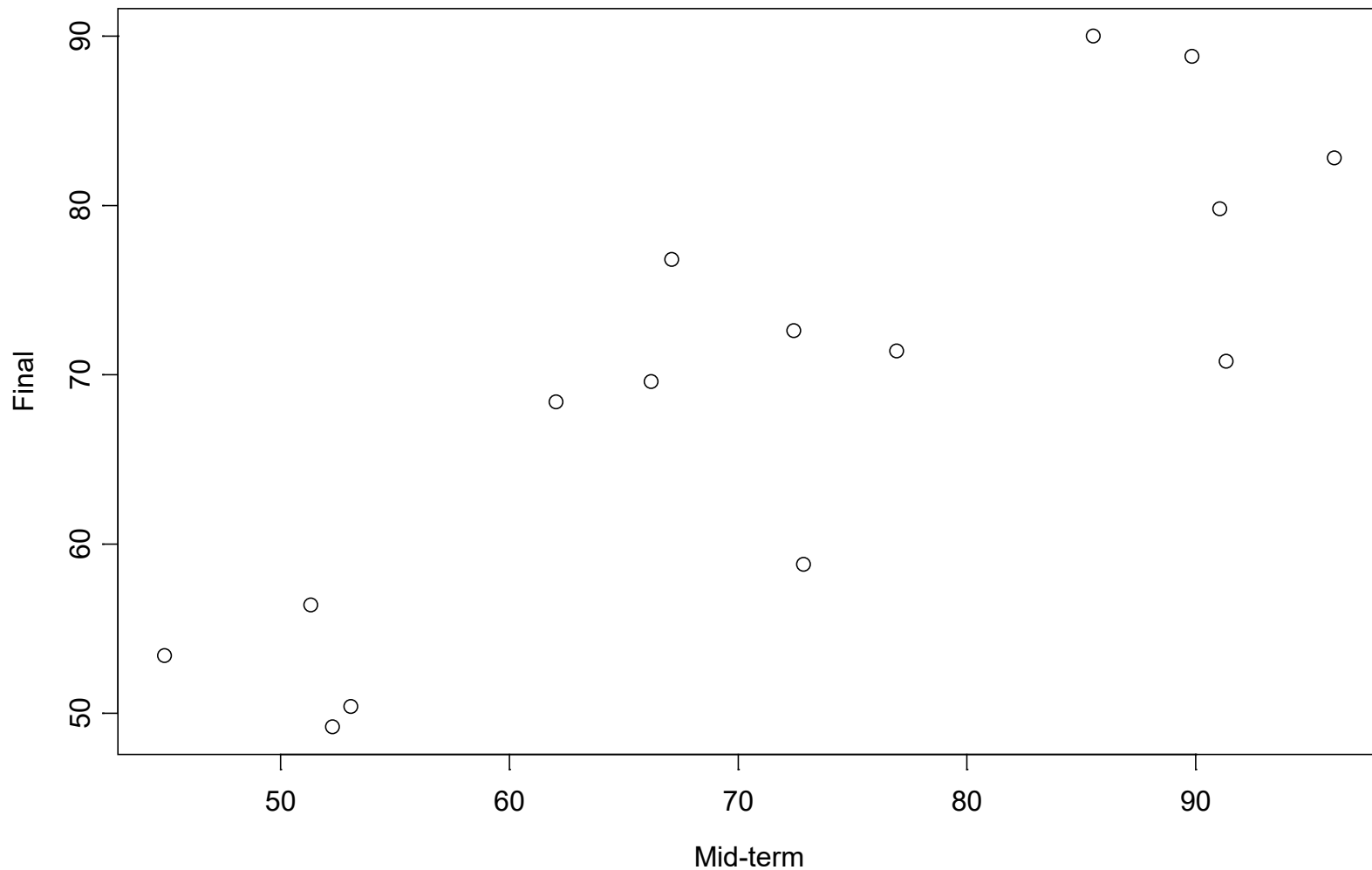
# Monte Carlo及Bootstrap的模擬程式

```
# Monte Carlo
t1=NULL
for (i in 1:10000) { x1=rpois(30,10); y1=median(x1); t1=c(t1,y1) }
# Bootstrap
x0=rpois(30,10)
t2=NULL
for (i in 1:10000) { x2=sample(x0,30,T); y2=median(x2); t2=c(t2,y2) }
#
a1=mean(t1)
b1=sqrt(var(t1))
b2=sqrt(var(t2))
par(mfrow=c(1,2))
hist(t1,xlab="Median",main="Monte Carlo (10,000 runs)")
text(8.2,3500,c("C.I.=(8.394, 11.277)"))
text(11.5,3500,c("s.e.=0.735"))
hist(t2,xlab="Median",main="Bootstrap (10,000 runs)")
text(8.7,3500,c("s.e.=1.006"))
text(12,3500,c("C.I.=(8.559, 11.441)"))
```

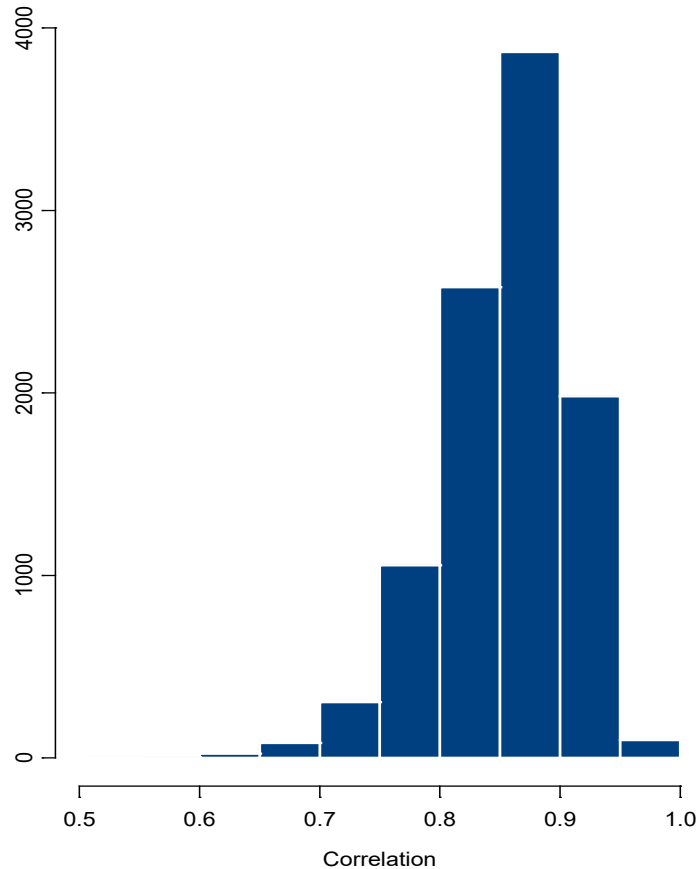
範例：75位選修統計學的學生，想瞭解期中考與總成績的相關性，只抽出15位學生。



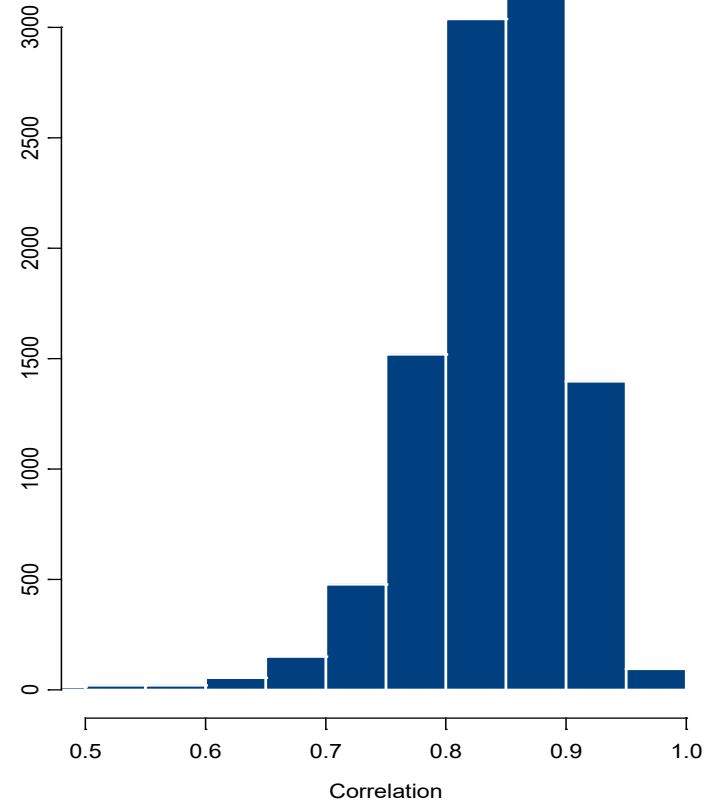
抽出的15位學生與母體特性大致接近，相關係數分別是0.8399及0.8543。



Monte Carlo



Bootstrap



比較各一萬次模擬，Monte Carlo法由母體任意抽出15個樣本得出的相關係數，與15個樣本以Bootstrap法算出的相關係數。

→ 兩者非常接近！(標準差0.0540 vs. 0.0654)



■ 註：

(1) 理論上，我們認為

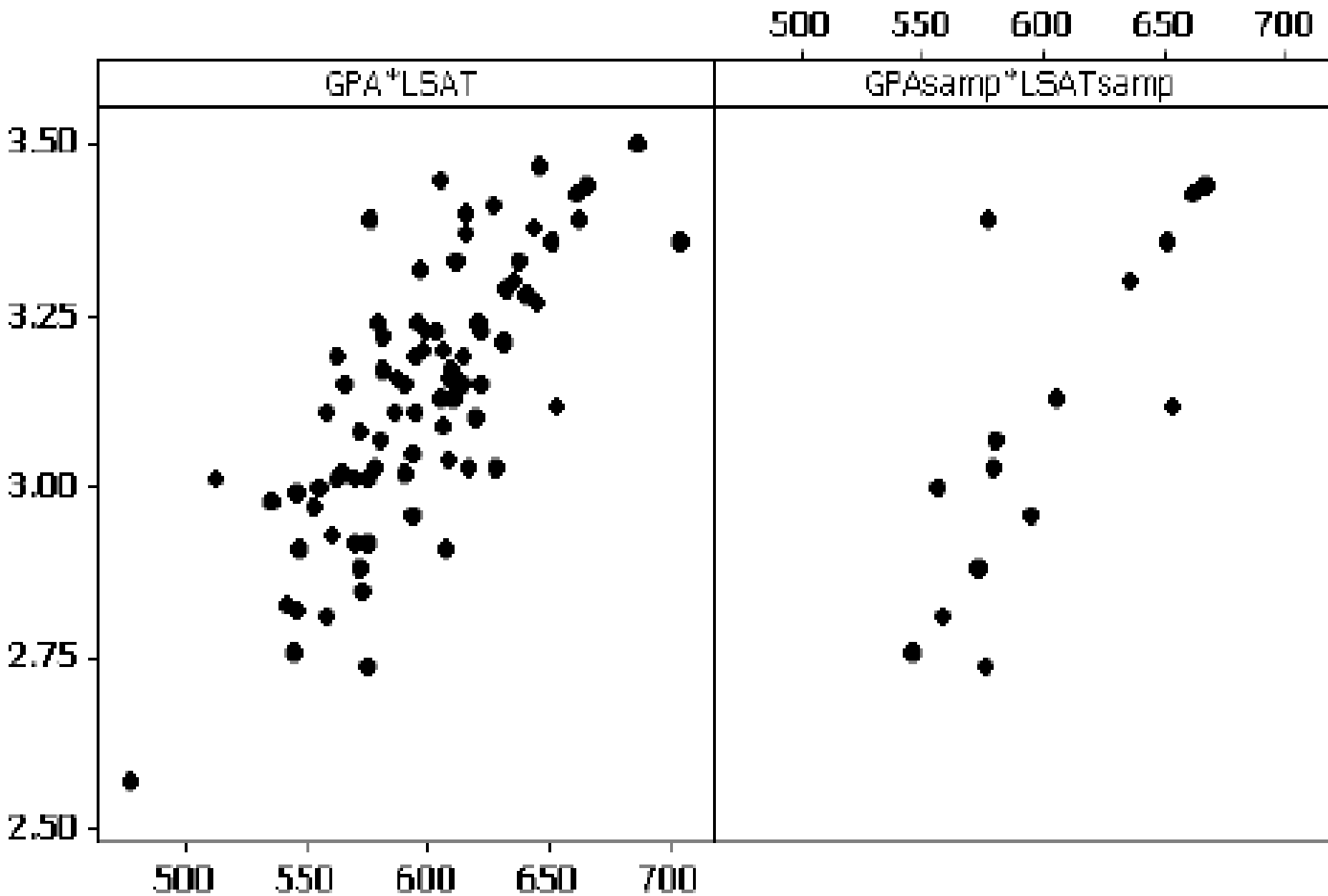
Bootstrap error

= Sampling error + bootstrap simulation error

(2) 當有充足的樣本數、且樣本具有與母體類似的特性時，Bootstrap可用來近似分配的形狀。(Shift methods!)

(3) 請參考Bootstrap講義！

# Law School Data Scatterplots: Population and Sample



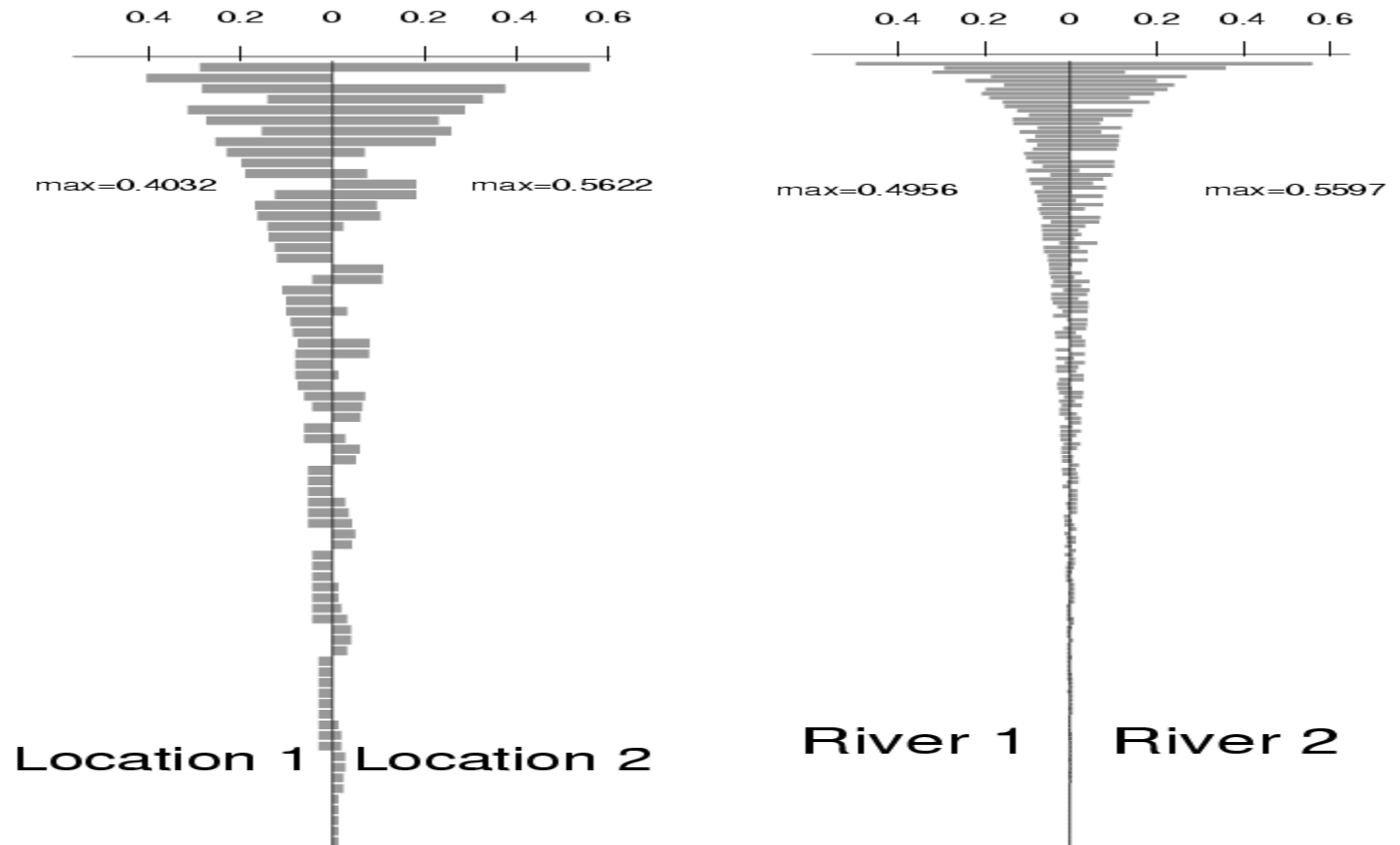


| school      | LSAT | GPA  | school      | LSAT | GPA  | school      | LSAT | GPA  | school      | LSAT | GPA  |
|-------------|------|------|-------------|------|------|-------------|------|------|-------------|------|------|
| 1           | 622  | 3.23 | 22          | 614  | 3.19 | 43          | 573  | 2.85 | 63          | 572  | 3.08 |
| 2           | 542  | 2.83 | 23          | 628  | 3.03 | 44          | 644  | 3.38 | 64          | 610  | 3.13 |
| 3           | 579  | 3.24 | 24          | 575  | 3.01 | <b>(45)</b> | 545  | 2.76 | 65          | 562  | 3.01 |
| <b>(4)</b>  | 653  | 3.12 | 25          | 662  | 3.39 | 46          | 645  | 3.27 | 66          | 635  | 3.30 |
| 5           | 606  | 3.09 | 26          | 627  | 3.41 | <b>(47)</b> | 651  | 3.36 | 67          | 614  | 3.15 |
| <b>(6)</b>  | 576  | 3.39 | 27          | 608  | 3.04 | 48          | 562  | 3.19 | 68          | 546  | 2.82 |
| 7           | 620  | 3.10 | 28          | 632  | 3.29 | 49          | 609  | 3.17 | 69          | 598  | 3.20 |
| 8           | 615  | 3.40 | 29          | 587  | 3.16 | <b>(50)</b> | 555  | 3.00 | <b>(70)</b> | 666  | 3.44 |
| 9           | 553  | 2.97 | 30          | 581  | 3.17 | 51          | 586  | 3.11 | 71          | 570  | 3.01 |
| 10          | 607  | 2.91 | <b>(31)</b> | 605  | 3.13 | <b>(52)</b> | 580  | 3.07 | 72          | 570  | 2.92 |
| 11          | 558  | 3.11 | 32          | 704  | 3.36 | <b>(53)</b> | 594  | 2.96 | 73          | 605  | 3.45 |
| 12          | 596  | 3.24 | 33          | 477  | 2.57 | 54          | 594  | 3.05 | 74          | 565  | 3.15 |
| <b>(13)</b> | 635  | 3.30 | 34          | 591  | 3.02 | 55          | 560  | 2.93 | 75          | 686  | 3.50 |
| 14          | 581  | 3.22 | <b>(35)</b> | 578  | 3.03 | 56          | 641  | 3.28 | 76          | 608  | 3.16 |
| <b>(15)</b> | 661  | 3.43 | <b>(36)</b> | 572  | 2.88 | 57          | 512  | 3.01 | 77          | 595  | 3.19 |
| 16          | 547  | 2.91 | 37          | 615  | 3.37 | 58          | 631  | 3.21 | 78          | 590  | 3.15 |
| 17          | 599  | 3.23 | 38          | 606  | 3.20 | 59          | 597  | 3.32 | <b>(79)</b> | 558  | 2.81 |
| 18          | 646  | 3.47 | 39          | 603  | 3.23 | 60          | 621  | 3.24 | 80          | 611  | 3.16 |
| 19          | 622  | 3.15 | 40          | 535  | 2.98 | 61          | 617  | 3.03 | 81          | 564  | 3.02 |
| 20          | 611  | 3.33 | 41          | 595  | 3.11 | 62          | 637  | 3.33 | <b>(82)</b> | 575  | 2.74 |
| 21          | 546  | 2.99 | 42          | 575  | 2.92 |             |      |      |             |      |      |

Sampled schools have bold-faced school numbers.

# Bootstrap法估計變異數的實例

- 檢定兩個數值是否相等。



- Bootstrap法計算出的變異數(標準差)，與大樣本理論Delta 法的數值比較：

|      | <b>Bootstrap</b> | <b>Delta</b> |
|------|------------------|--------------|
| 螃蟹資料 | 0.015            | 0.01426      |
| 水鳥資料 | 0.008            | 0.0083       |



# Bootstrap用於檢定與預測

- 除了用於估計變異數外，也可用於檢定與預測。
  - 檢定可與變異數估計結合，在此將略過細節，僅以一個範例說明想法。
  - 預測多半與相關資料(Dependent data)有關，都屬於無母數方法，但至今仍無統一的方法。常見的方法有Block, Sieve, Local, Wild與Markov Bootstrap，以及Subsampling，在此只介紹Block Bootstrap。

# Bootstrap法的檢定實例

■ 檢定某個假設是否成立。

→ 範例：死亡率是否服從Gompertz 分配  
也就是瞬間死亡率滿足

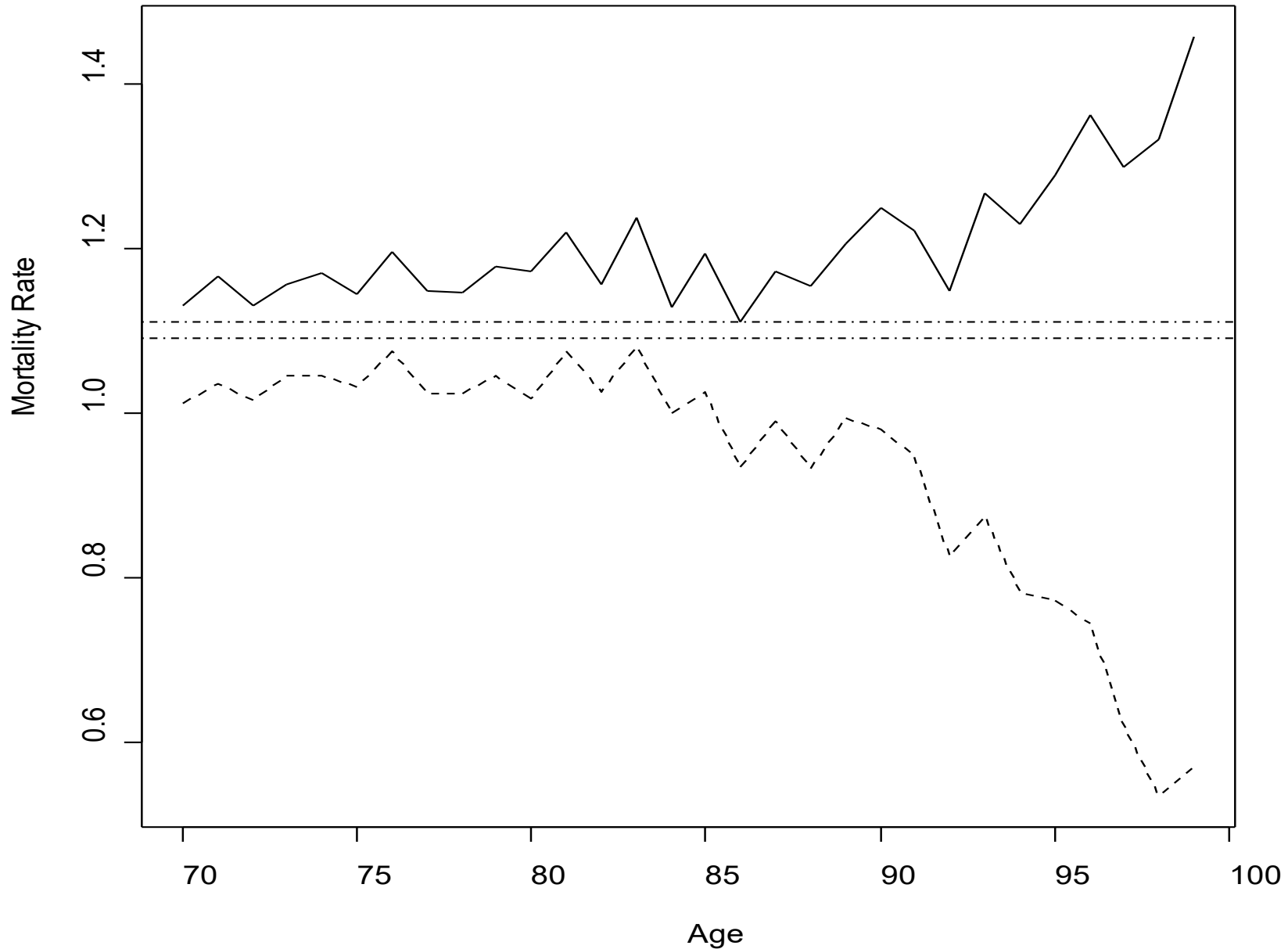
$$\mu_x = BC^x, \quad B > 0, C > 1$$

或是

$$\log(p_{x+1})/\log(p_x) = C$$

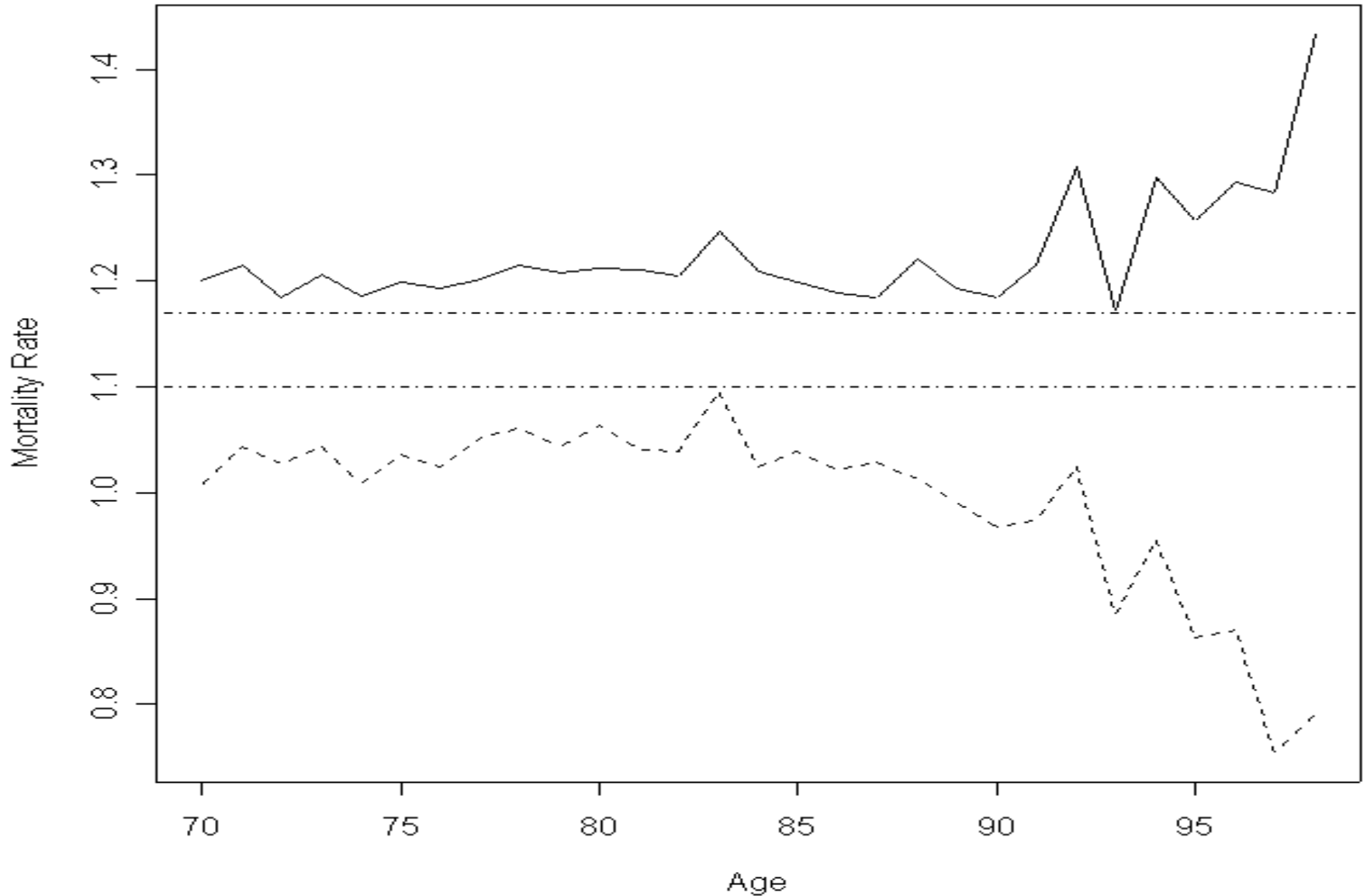
若C的可能分布範圍不為空集合，則不拒絕Gompertz假設。

# Bootstrap C.I. of Male Mortality Rat



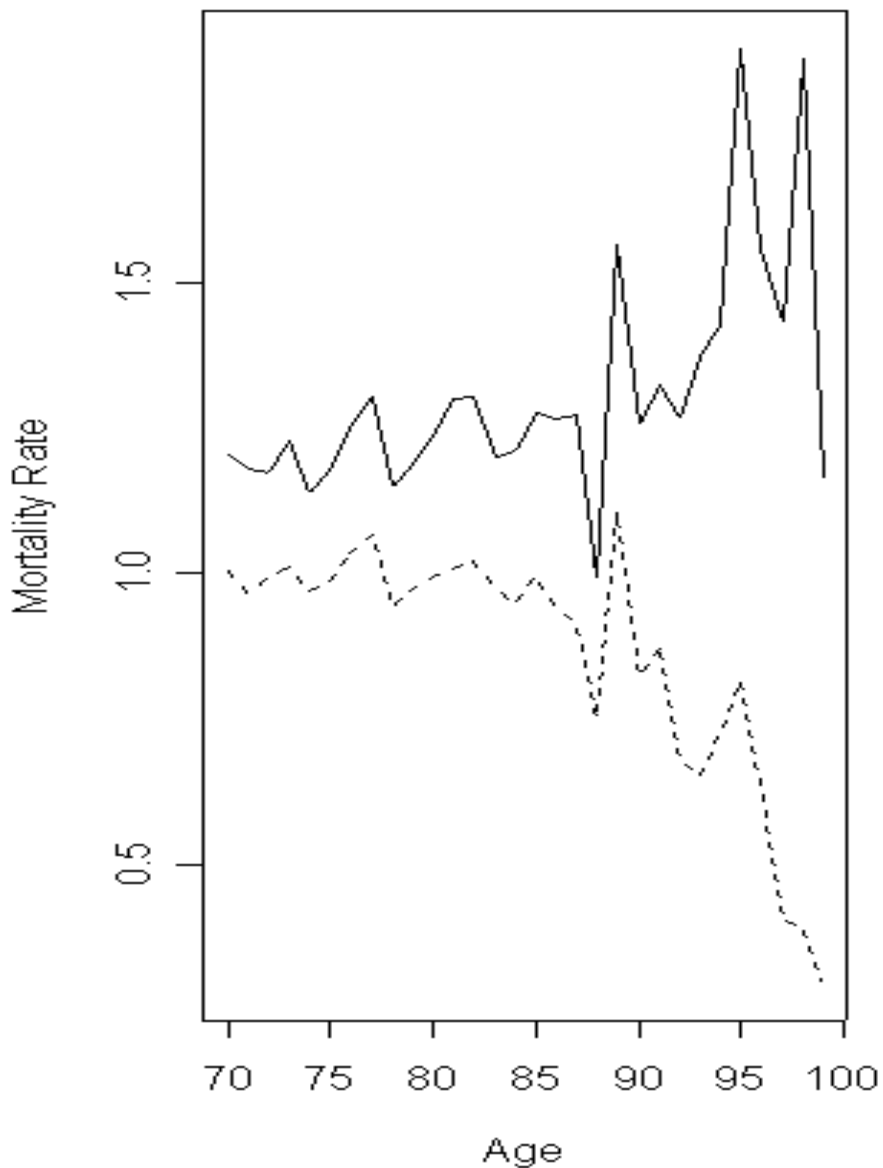
台灣男性死亡率1000次Bootstrap模擬(1999~2001年)

## Bootstrap C.I. of Female Mortality Rates

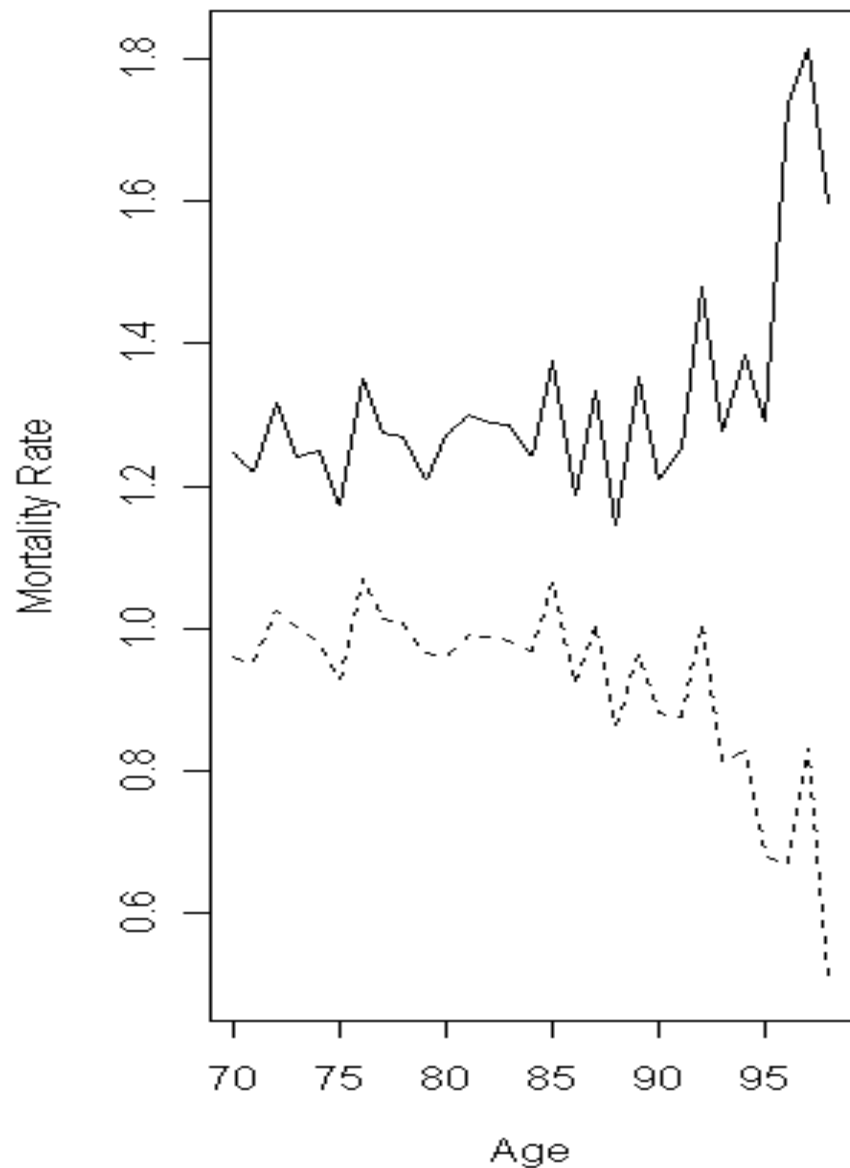


台灣女性死亡率1000次Bootstrap模擬(1999~2001年)

**C.I. for Gompertz (2001 Male)**



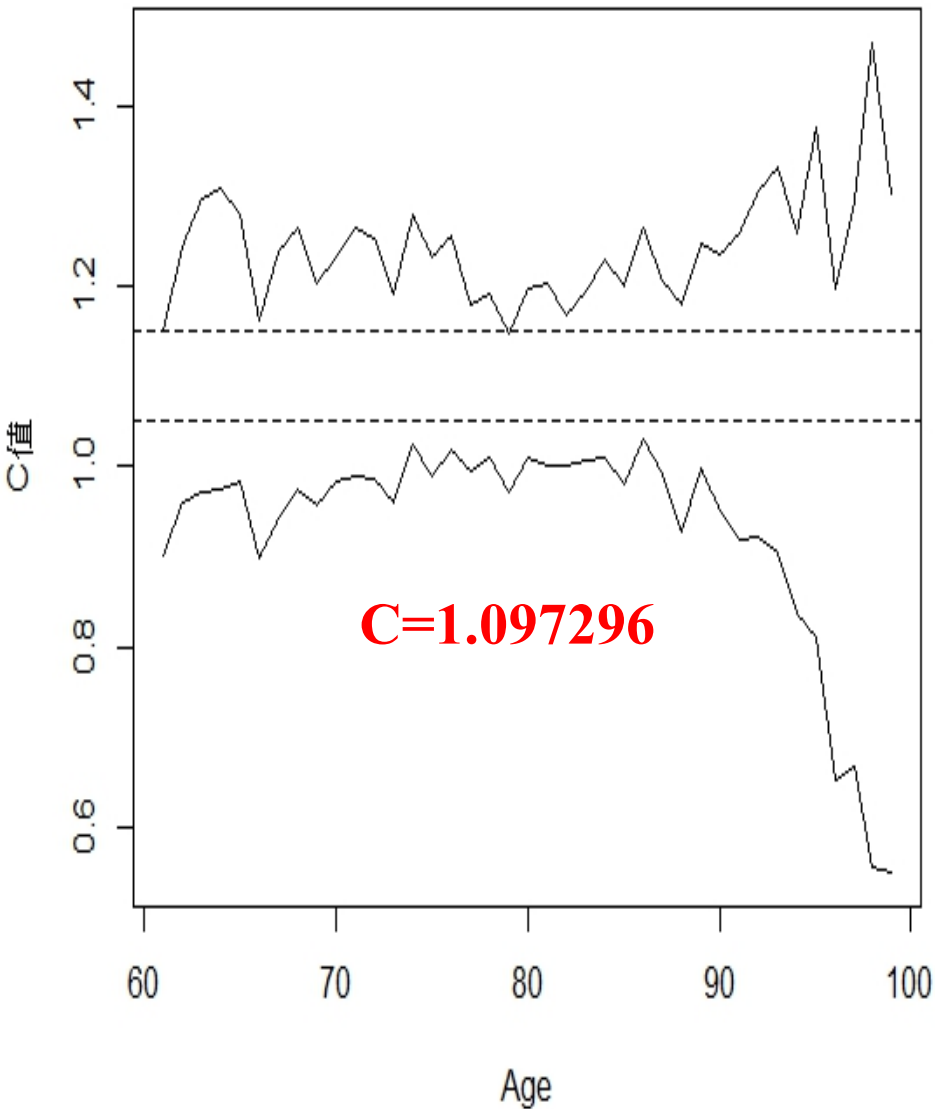
**C.I. for Gompertz (2001 Female)**



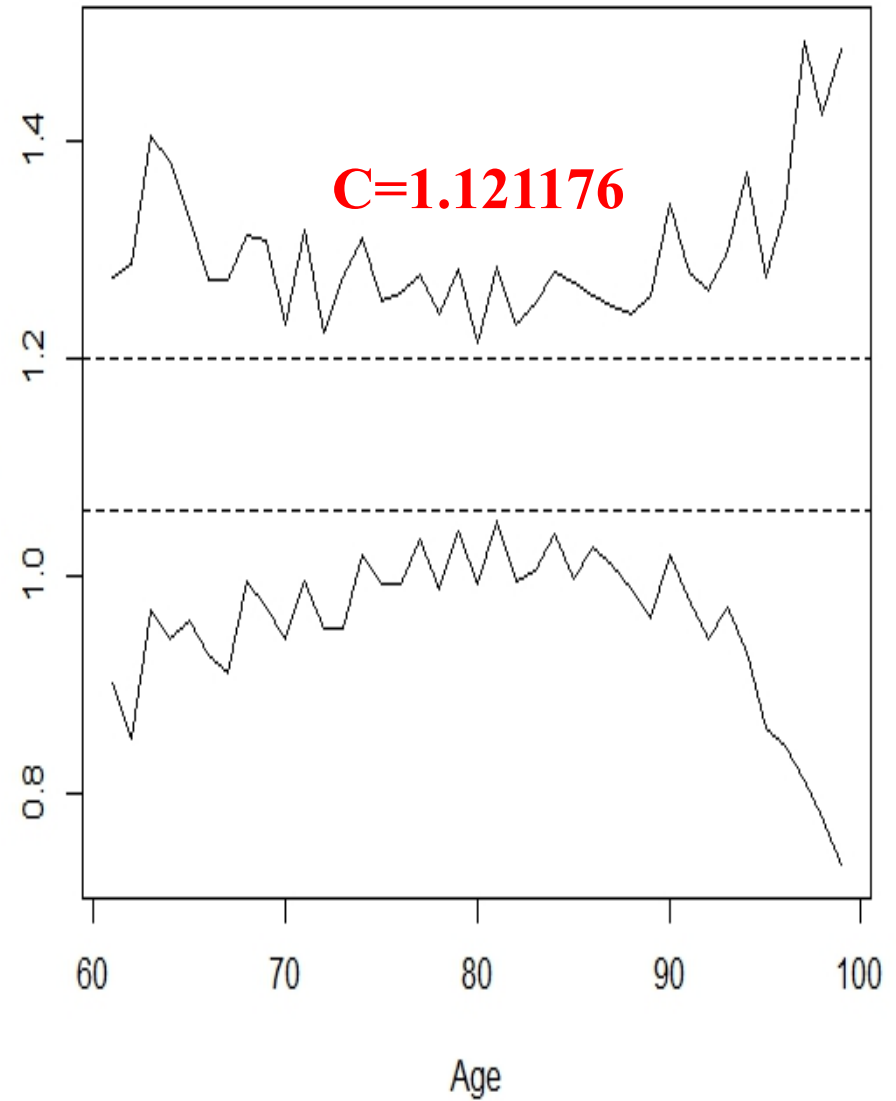
台灣男女兩性死亡率1000次Bootstrap模擬(2001年)



C value for men



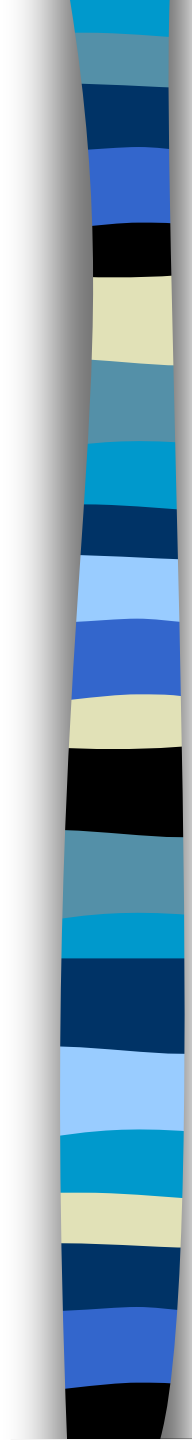
C value for Women



2009至2011年臺灣高齡死亡率的模擬檢定

# 區塊(Block)拔靴法

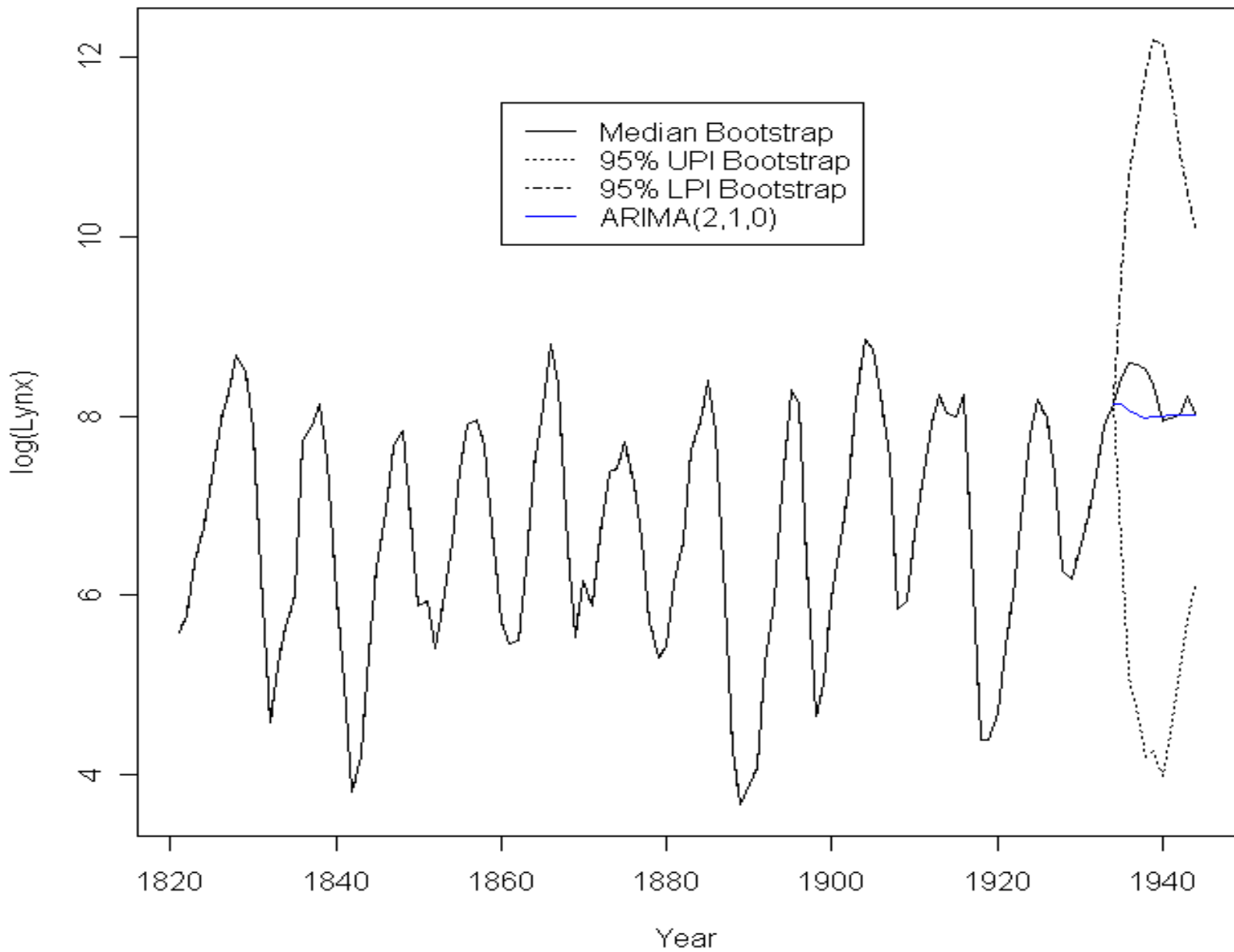
- 區塊拔靴法的重複抽樣方式類似一般的拔靴法，只是每次抽取一個「區塊」的資料，當資料服從均衡(Stationary)假設時，區塊拔靴法大多都適用。
- 區塊拔靴法用於相關資料有不錯的效果，雖然不如獨立樣本時一般拔靴法的準確，但比Subsampling效果好，而且不需要資料滿足很強的條件，加上操作時不需對資料給予任何假設，實證上是很好的選擇。

- 
- 區塊拔靴法的抽樣方式類似拔靴法，但不是對觀察值直接抽樣，而是對相鄰觀察值的差異抽樣，而且抽取時將連續一串的差異值抽出。例如：若區塊長度為 $b$ 、且抽到第 $k$ 個觀察值，則第 $k$ 個至第 $k + b - 1$ 個差異值被抽出，最後一個觀察值加上這些差異值即為預測值。
  - R的模組「boot」也有處理時間數列資料(相關資料的一種)的功能，細節可查閱「tsboot」指令。(這個指令可指定區塊長度為定值、或是服從幾何分配。)

- 範例：R的說明檔中使用「lynx」資料，一共有114筆資料。使用tsboot指令，可限定固定區塊長度，也可指定區塊長度服從幾何分配。（兩者得出的結果非常接近。）

→ 首先對「lynx」資料取對數，再計算相鄰時間觀察值的差值，之後抽出一整個區塊差值，加到最後一期的觀察值。下圖為依據一千次區塊拔靴法的模擬，計算而得的95%預測區間，與差分後的AR(2)模型之預測值之比較。

# Block Bootstrap in Lynx Data



# 以區塊拔靴法推估生育率、死亡率等數值：

差分後隨機  
抽取第[t~t+n]  
年的變動量

|       | 0~4歲        | 5~9歲        | ... | ... |
|-------|-------------|-------------|-----|-----|
| 1980年 | $q_x(1980)$ | $q_x(1980)$ | ... | ... |
| ⋮     | ⋮           | ⋮           | ⋮   | ⋮   |
| 2015年 | $q_x(2015)$ | $q_x(2015)$ | ... | ... |

+

|      |                   |                   |     |     |
|------|-------------------|-------------------|-----|-----|
| t年   | $\Delta q_x(t)$   | $\Delta q_x(t)$   | ... | ... |
| ⋮    | ⋮                 | ⋮                 | ⋮   | ⋮   |
| t+n年 | $\Delta q_x(t+n)$ | $\Delta q_x(t+n)$ | ... | ... |

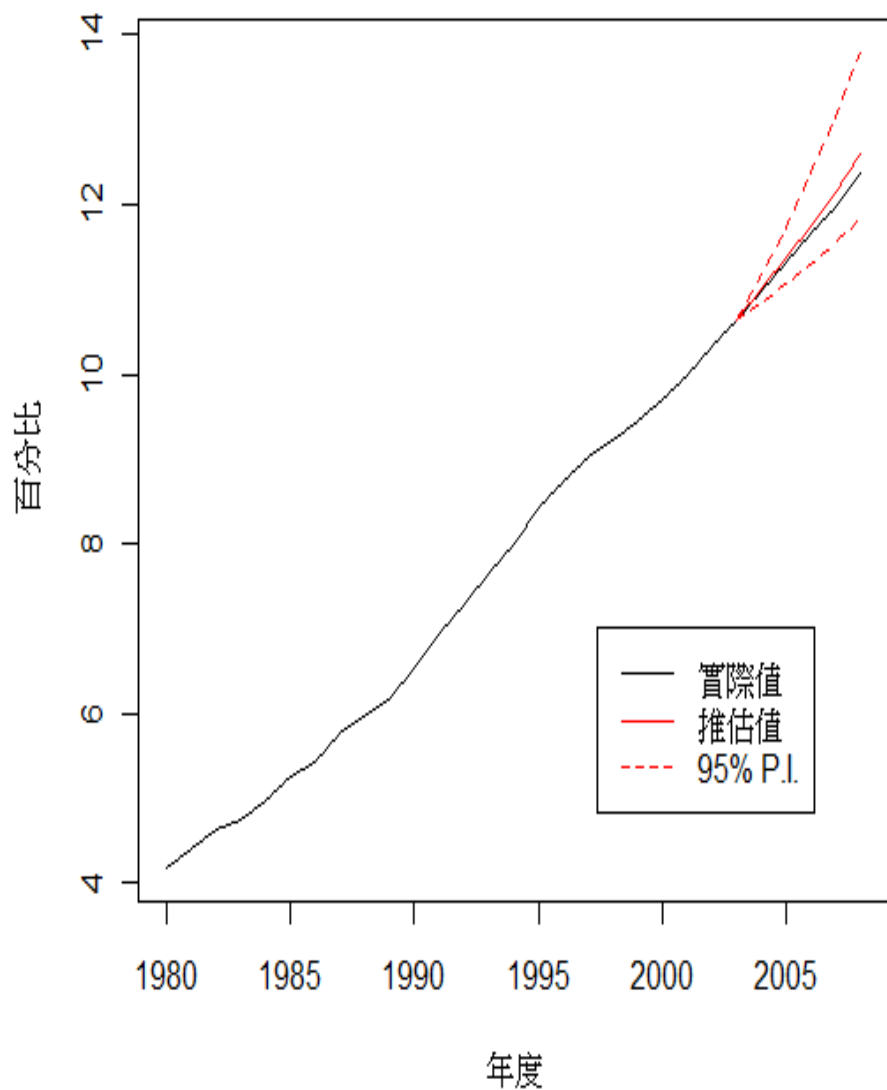
||

資料最後一年  
加上變動量，  
得到未來n年  
的推估結果

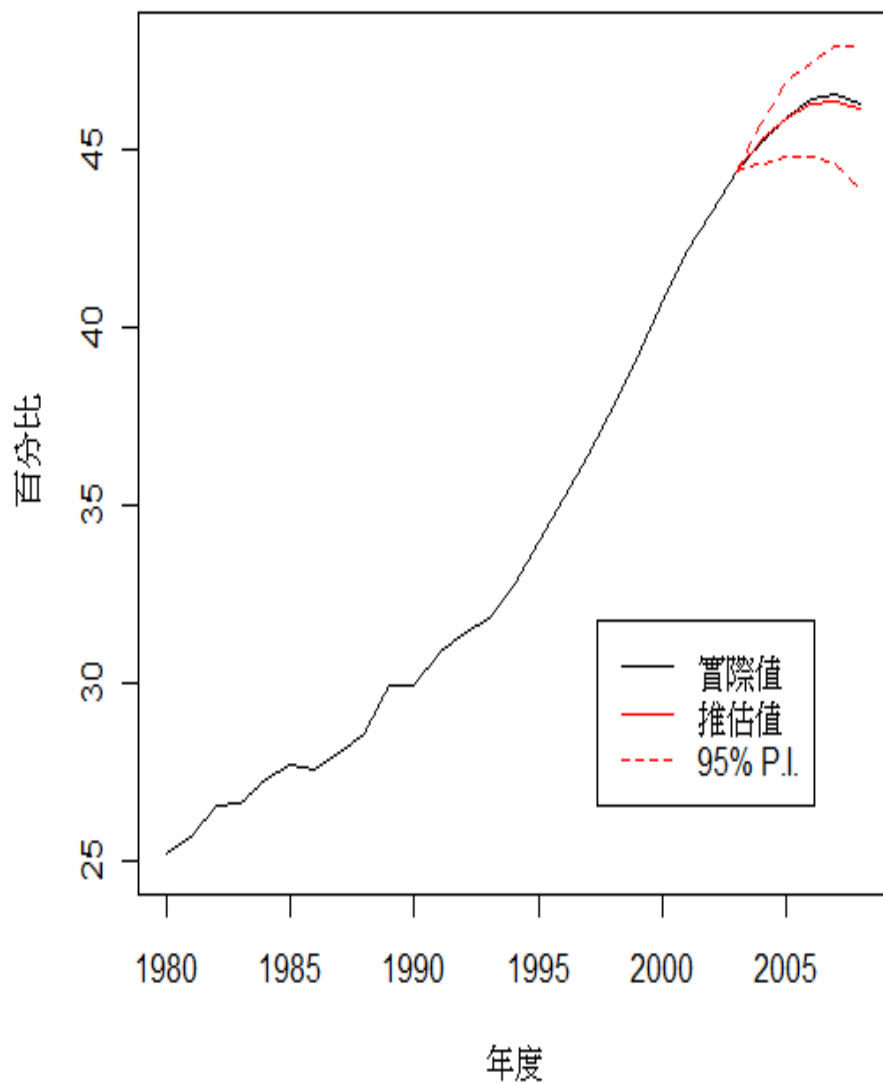
|       | 0~4歲        | 5~9歲        | ... | ... |
|-------|-------------|-------------|-----|-----|
| 2016年 | $q_x(2016)$ | $q_x(2016)$ | ... | ... |
| ⋮     | ⋮           | ⋮           | ⋮   | ⋮   |

# 區塊拔靴法範例：臺北市高齡人口比例

## 65 到 99 歲人口比例



## 75 到 99 歲人口佔 65 到 99 歲人口比例



# 臺北市主計處與區塊拔靴法 推估誤差比較

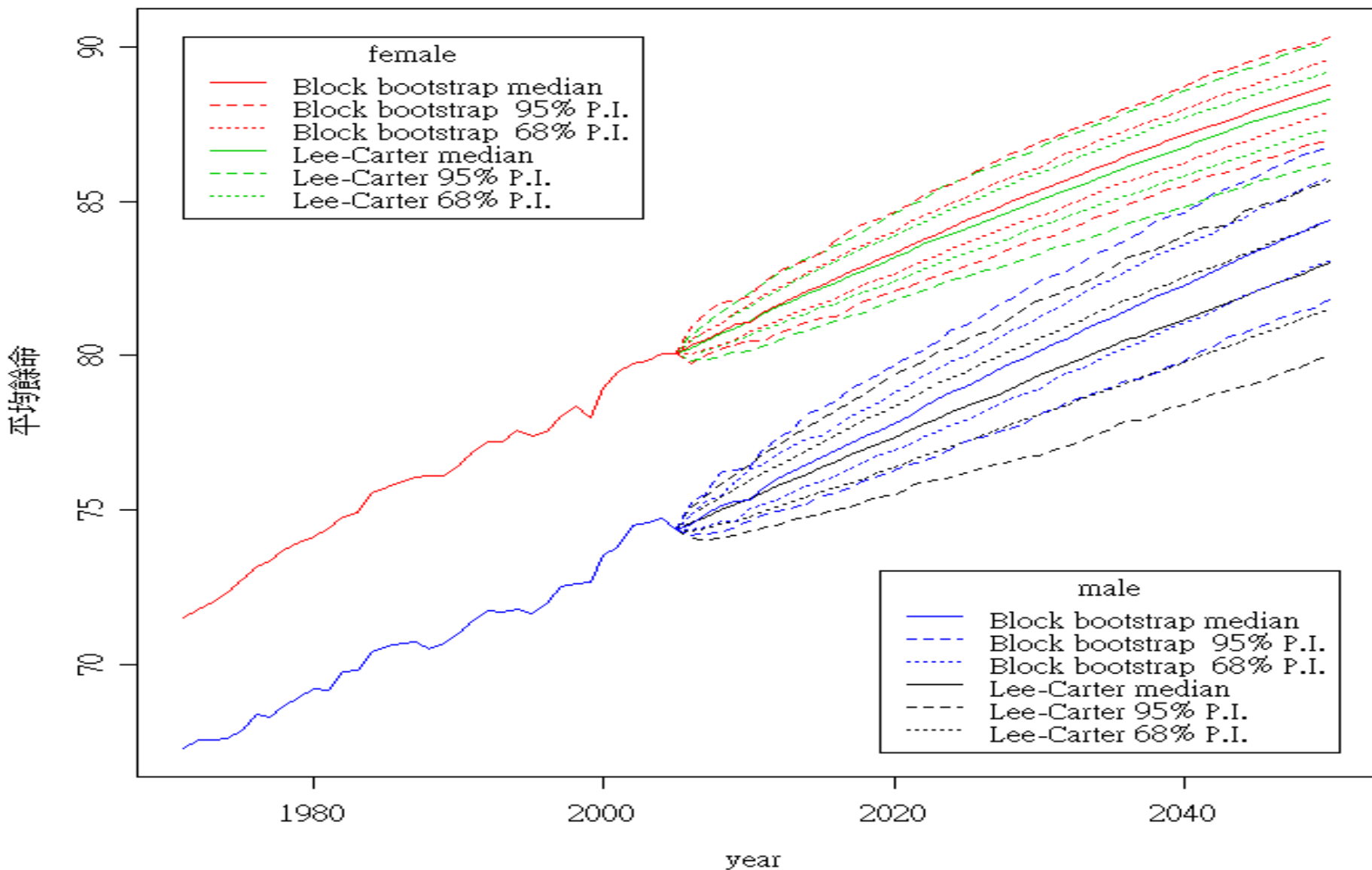
單位：%

| 要素  | 本研究  | 臺北市主計處 |       |               |               |
|-----|------|--------|-------|---------------|---------------|
|     |      | 世代生存法  | 迴歸分析法 | ARIMA<br>(直接) | ARIMA<br>(間接) |
| 總人口 | 0.80 | 1.70   | 3.17  | 0.87          | 1.38          |



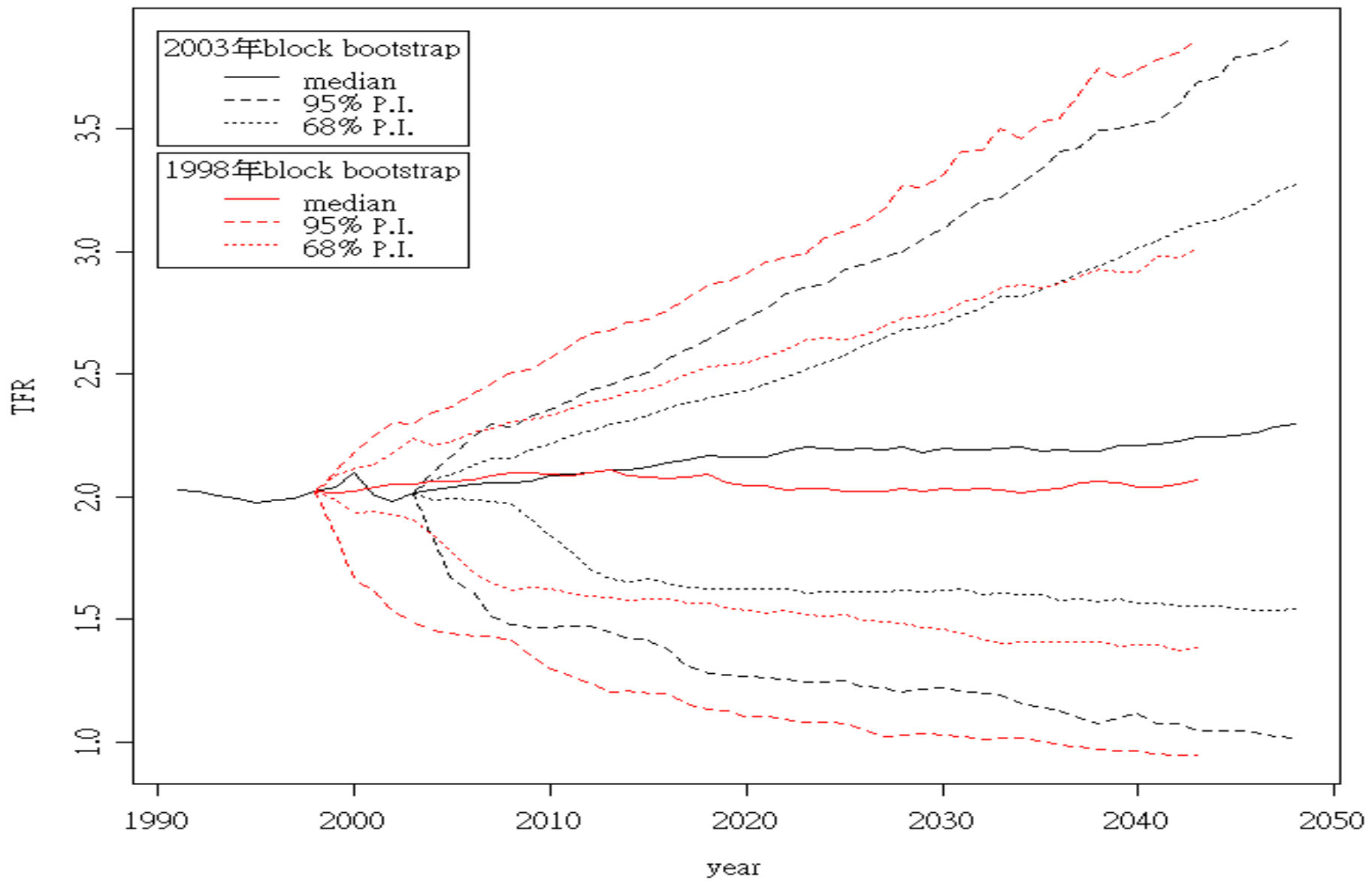
# 區塊拔靴法範例：臺灣死亡率模型預測

TAIWAN 零歲平均餘命

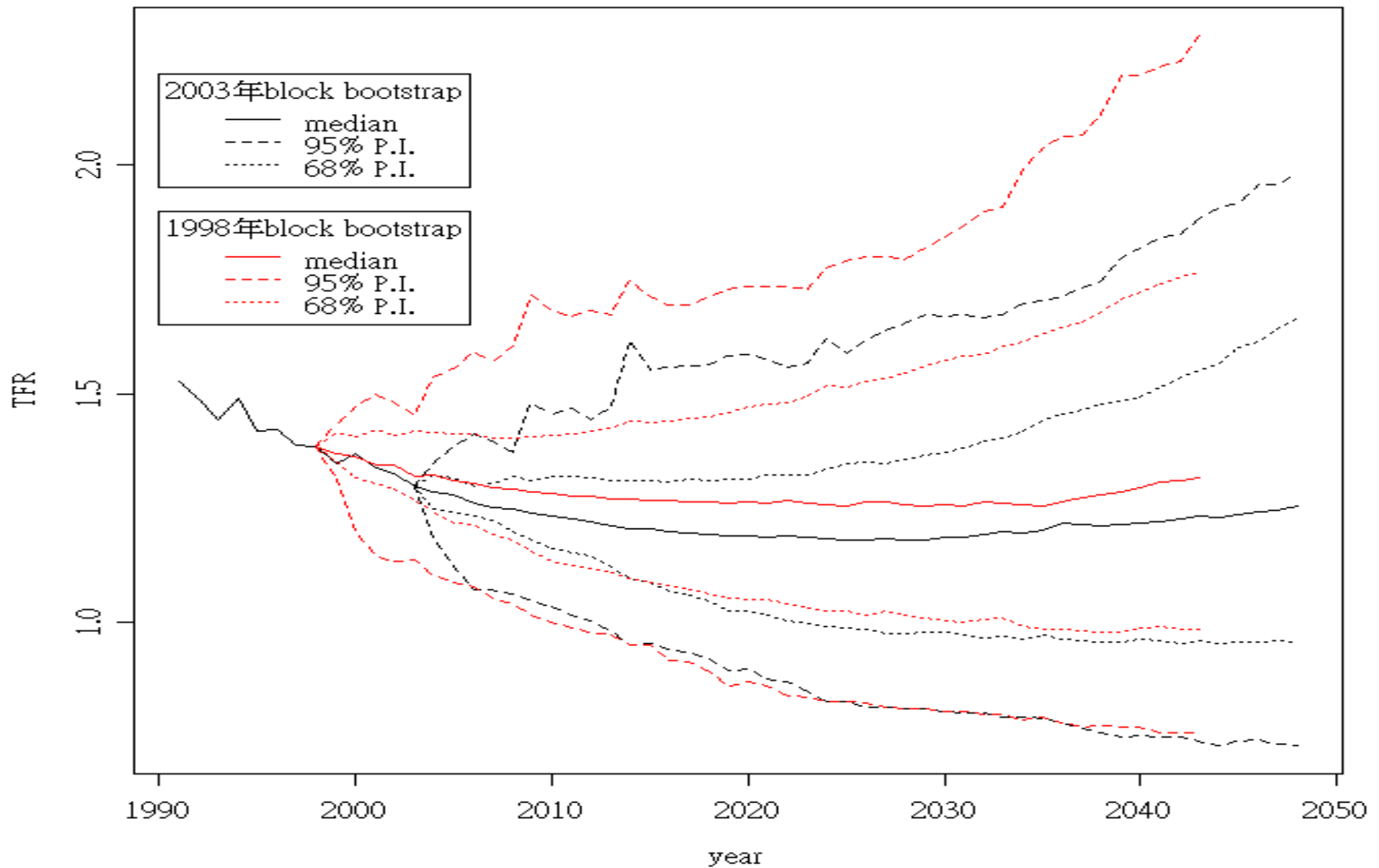


# 區塊拔靴法範例：美國總生育率預測

USA TFR

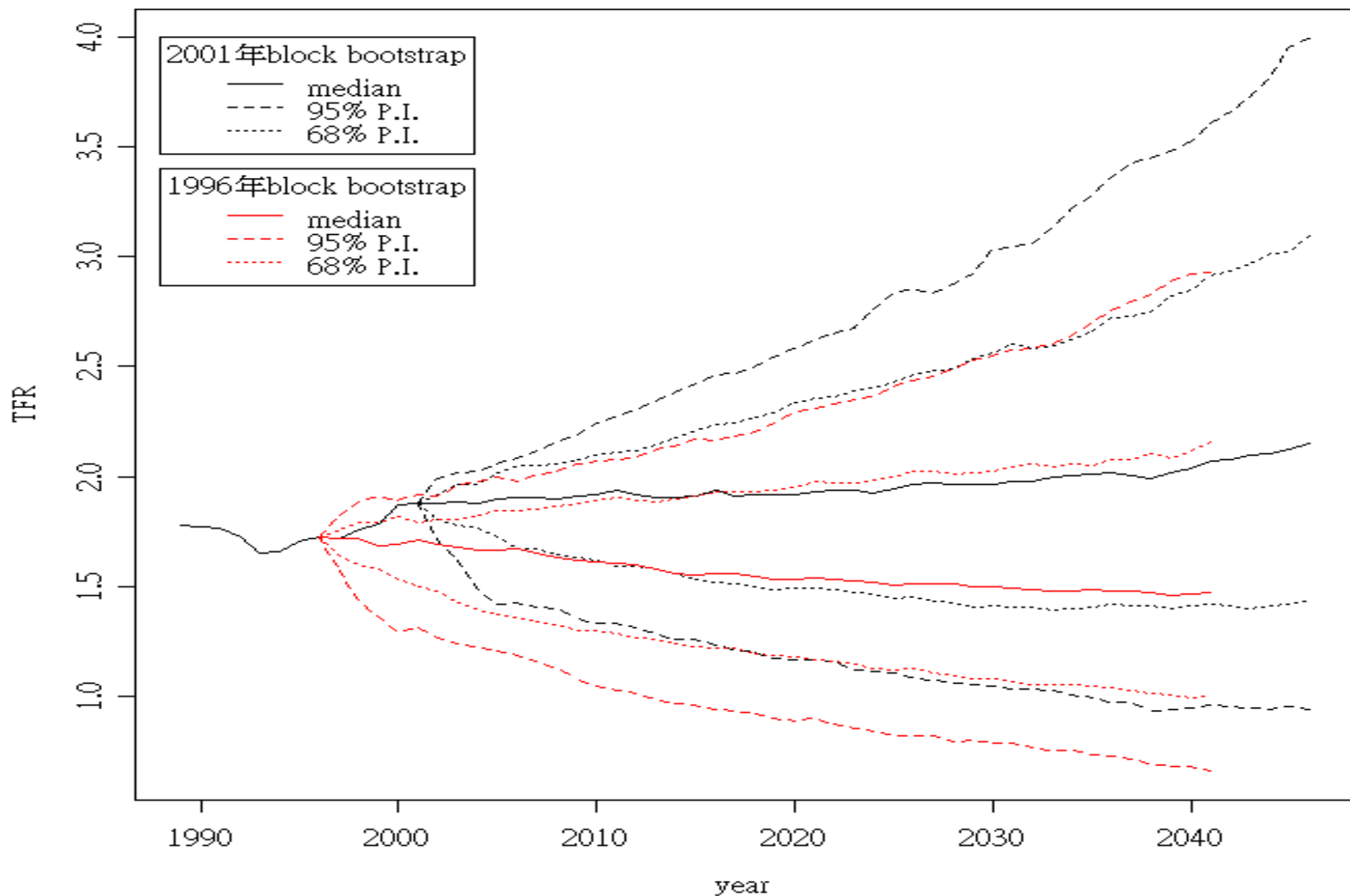


# JAPAN TFR



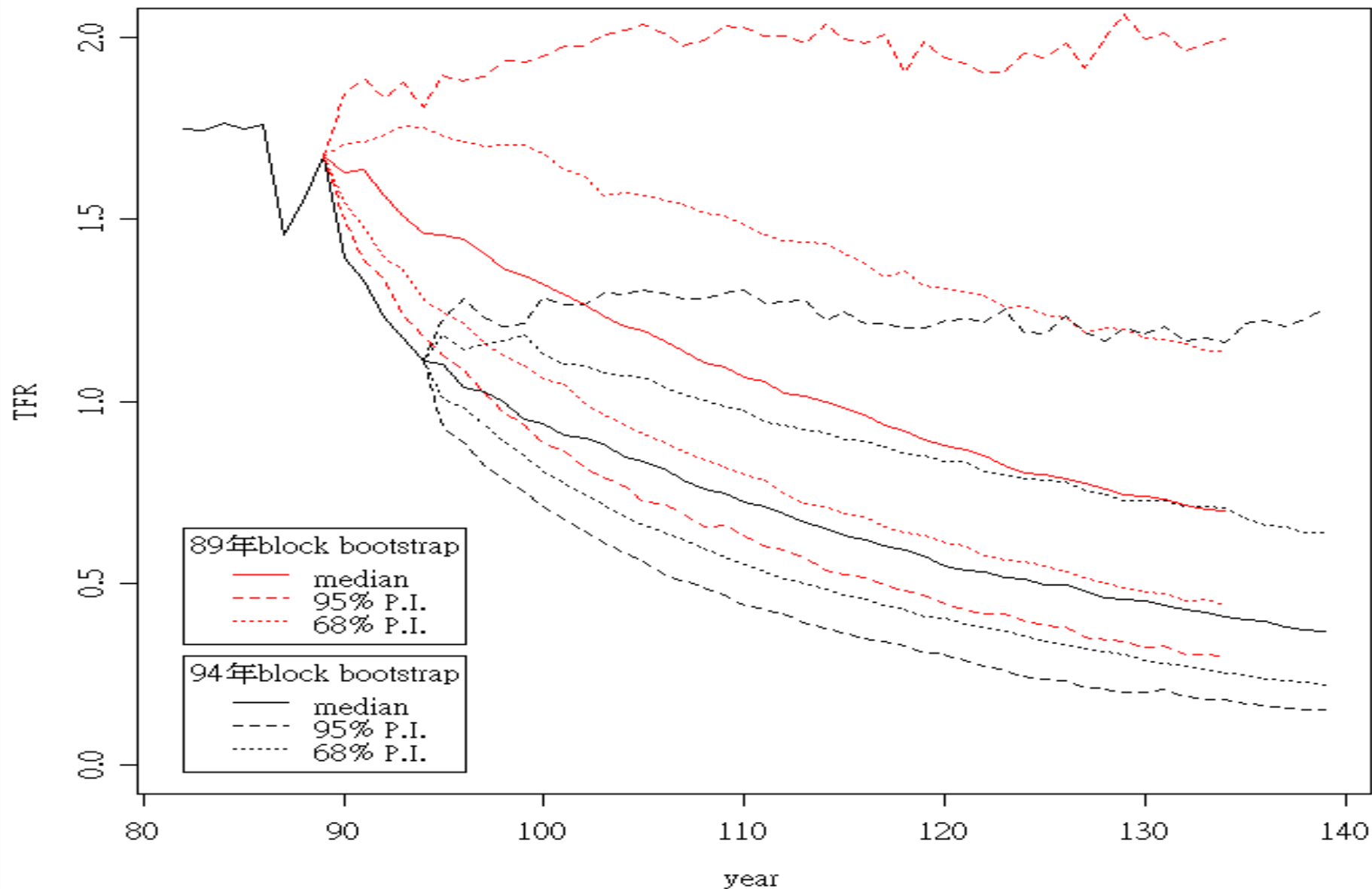
區塊拔靴法的交叉驗證(日本生育率)

# FRANCE TFR



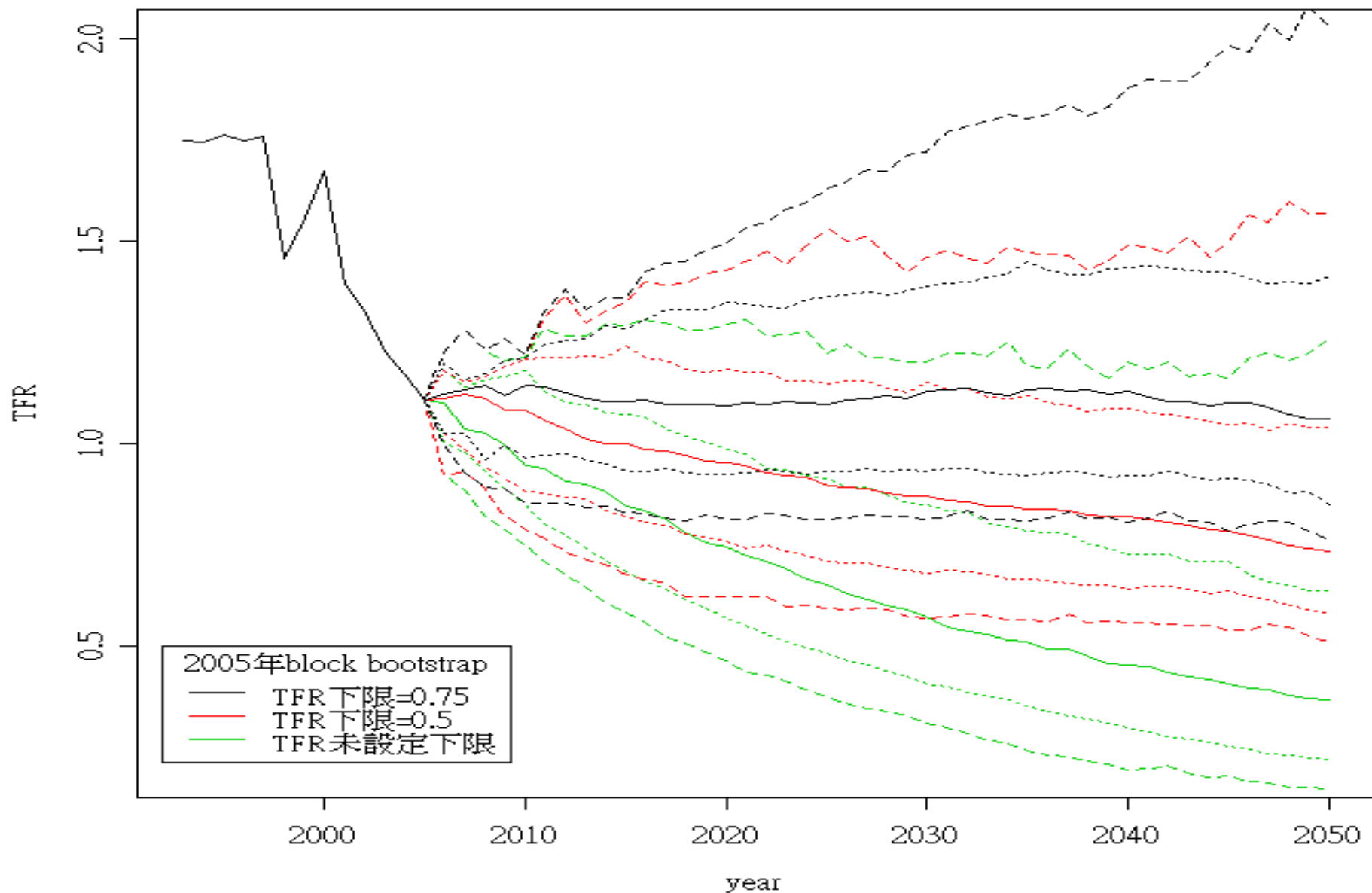
區塊拔靴法的交叉驗證(法國生育率)

# TAIWAN TFR



區塊拔靴法的交叉驗證(台灣生育率)

# TAIWAN TFR



設定下限的區塊拔靴法(台灣生育率)

# 股市投資策略模擬

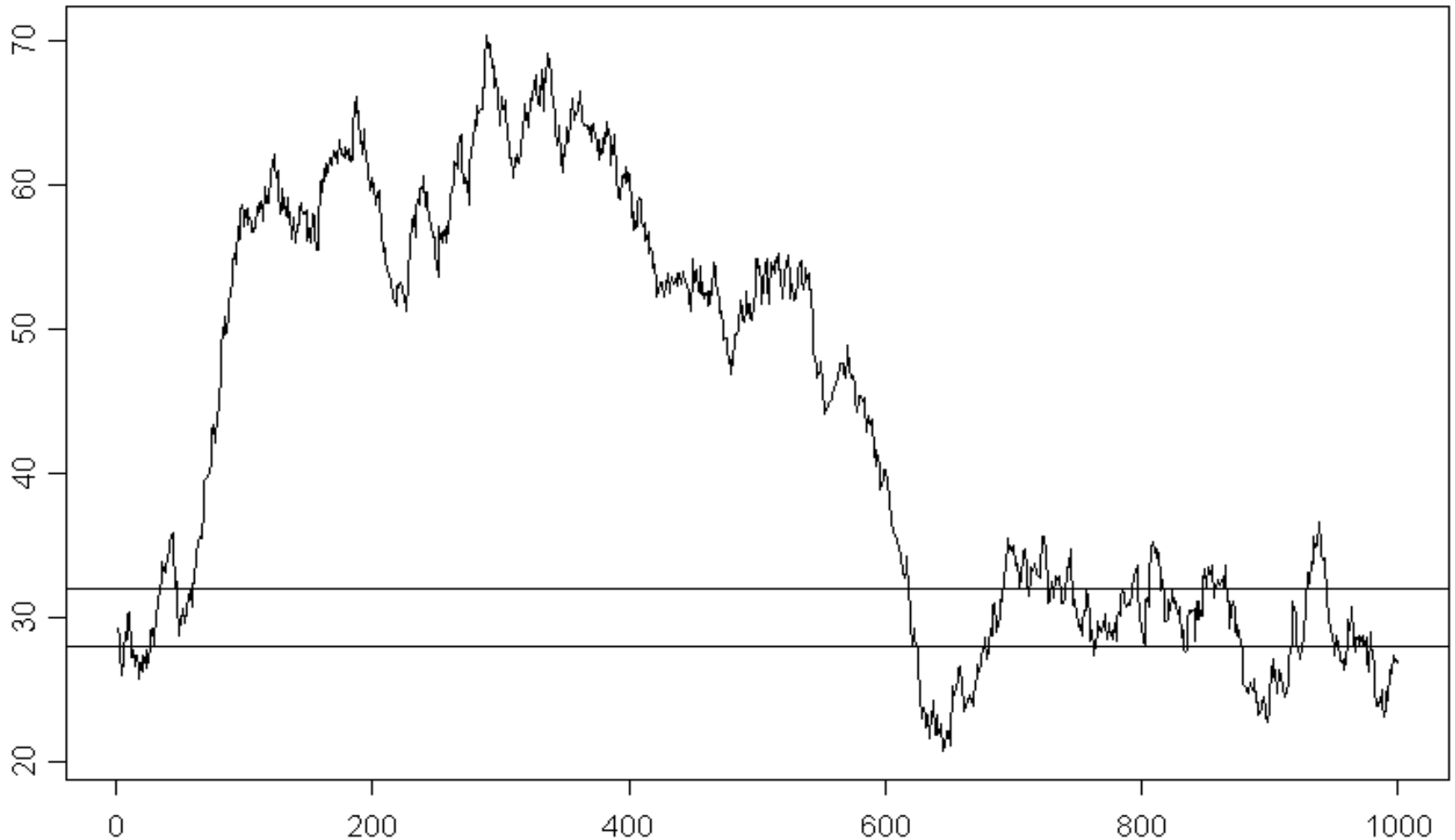
- 亞洲股市中據說台灣股市最接近隨機漫步(Random Walk)，如何在雜亂無章的訊息中獲取利潤？

→ 隨機漫步也就是假設時間  $t$  的股價  $B(t)$ ：

$$B(t+s) - B(t) | B(t) \sim N(s\mu, s\sigma^2)$$

→ 「逢低買進、逢高賣出！」

Time Series Plot of Simulated Stock Price



1000次模擬的年平均報酬為16%！



# 排隊理論的等待時間模擬

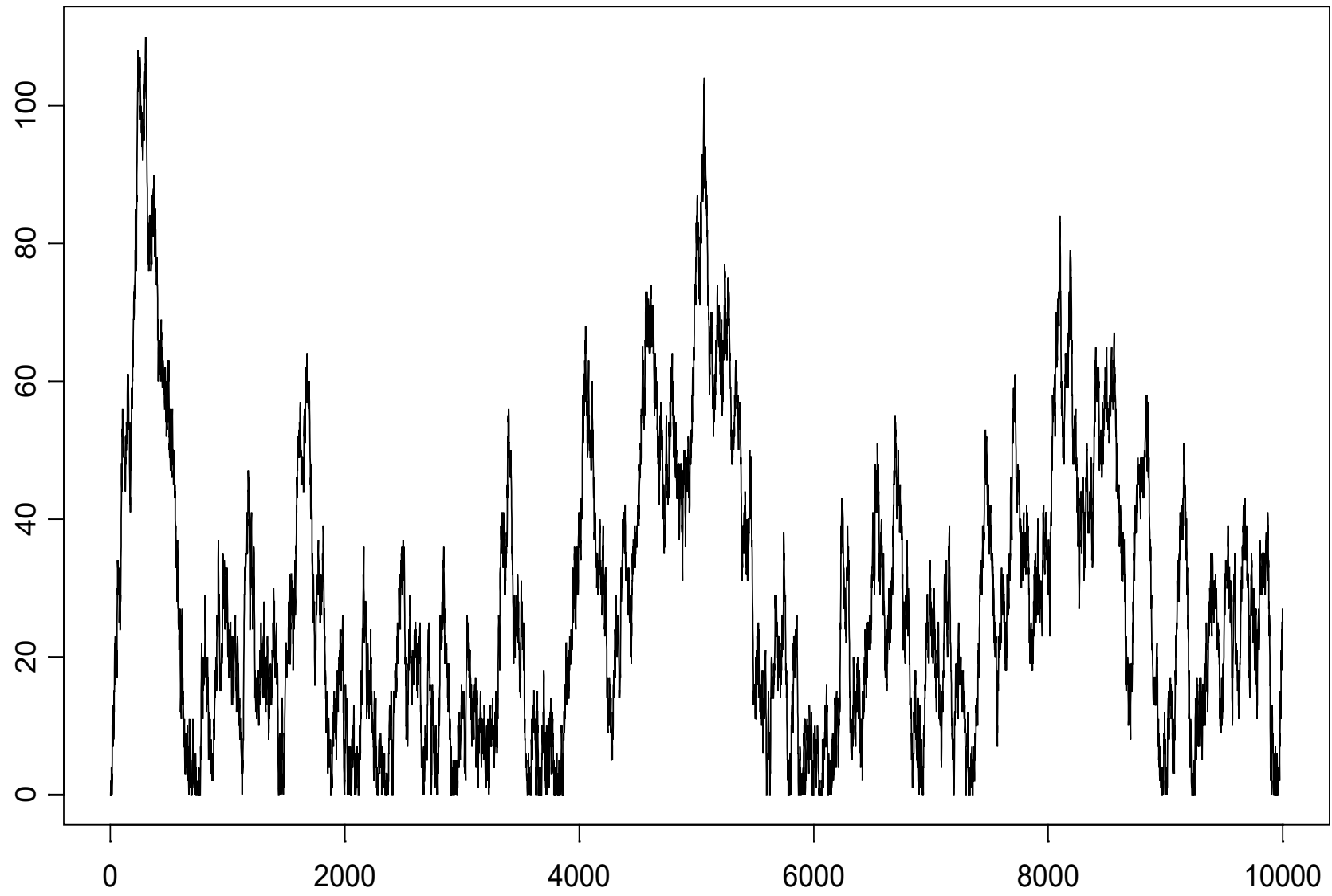
- 若服務每位顧客平均需時  $\mu$  分鐘的指數分配，兩位顧客間的時間為  $\lambda$  分鐘的指數分配；若  $\mu < \lambda$  則等待服務的顧客隊伍才有可能消失，否則將愈排愈長。

兩組模擬比較：(利用率 =  $\mu/\lambda$ )

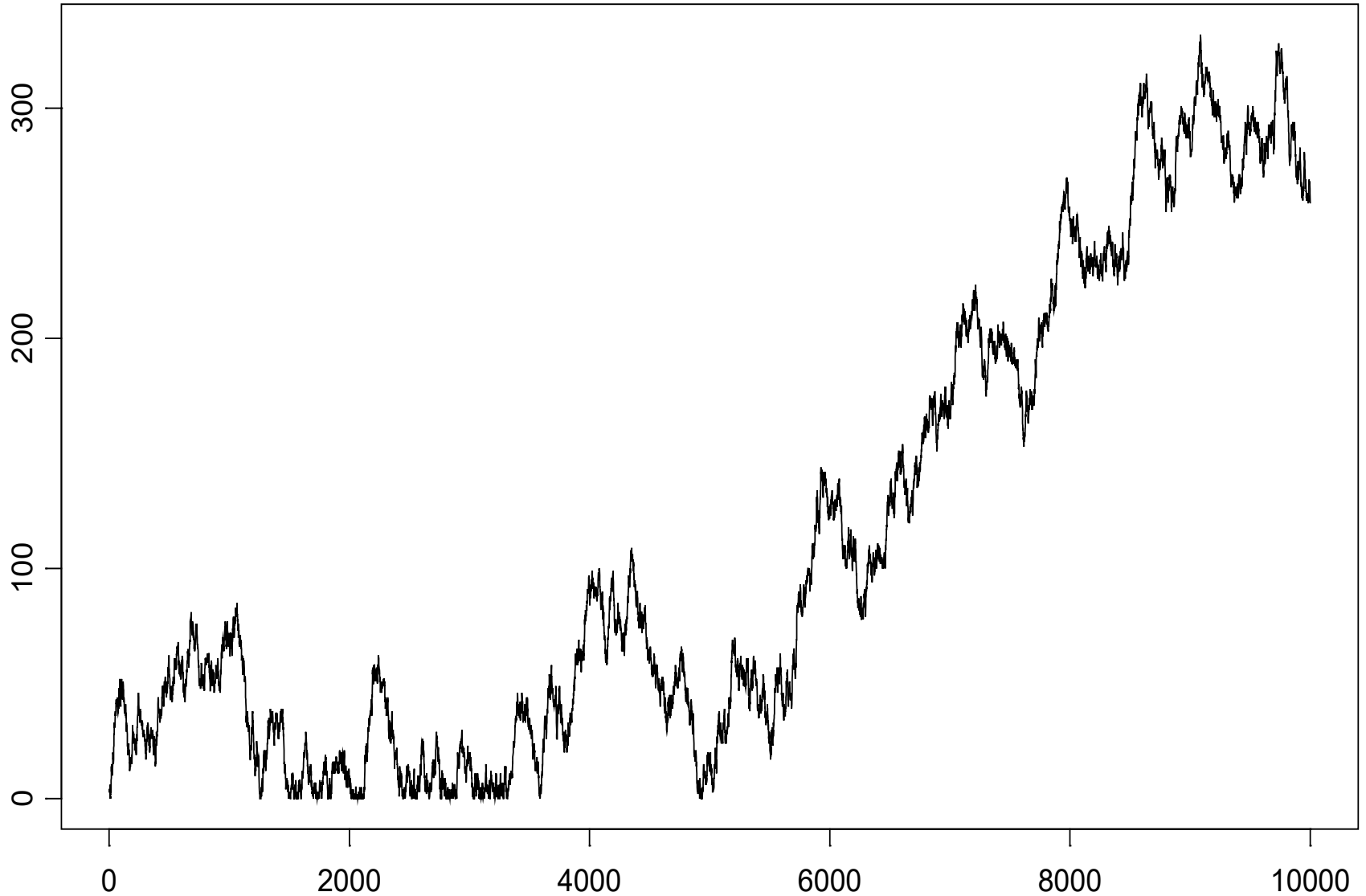
→  $\mu = 3.26 < \lambda = 3.27$ ；

→  $\mu = 3.26 > \lambda = 3.25$ 。

Time series plot when the interarrival time is bigger (3.27 v.s. 3.26)



Time series plot when the inter-arrival rate is smaller(3.25 v.s. 3.26)

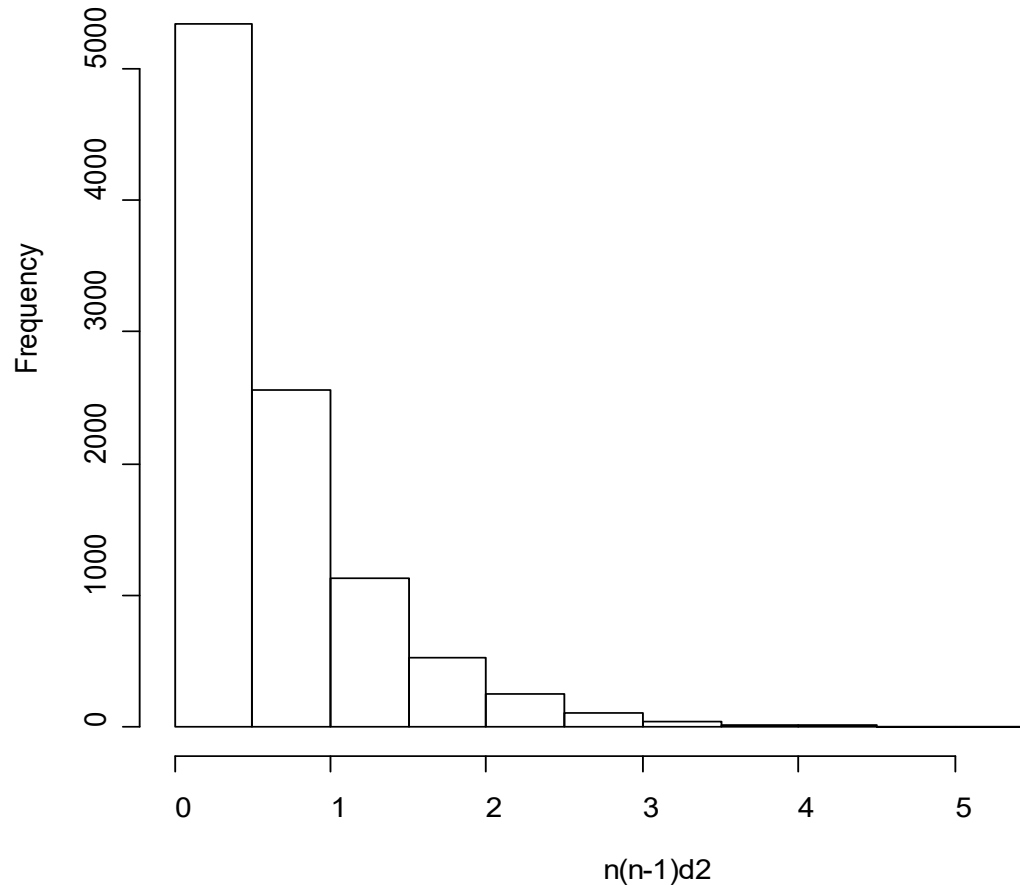


# 測試大樣本理論

- 當大樣本理論不容易推導(或不容易推敲)時，可用蒙地卡羅法測試樣本統計量的可能性質。
- 例如：在單位正方形 $(0,1) \times (0,1)$ 內隨機抽出  $n$  個觀察值，這些觀察值間的最小距離為  $d$ ，據說  $n(n-1)d^2$  服從某種分配，我們可以由電腦模擬探索這個分配開始。

下圖為  $n = 20$  的一萬次模擬所得出的長條圖，似乎與指數分配有關：

Histogram of 10,000 simulation

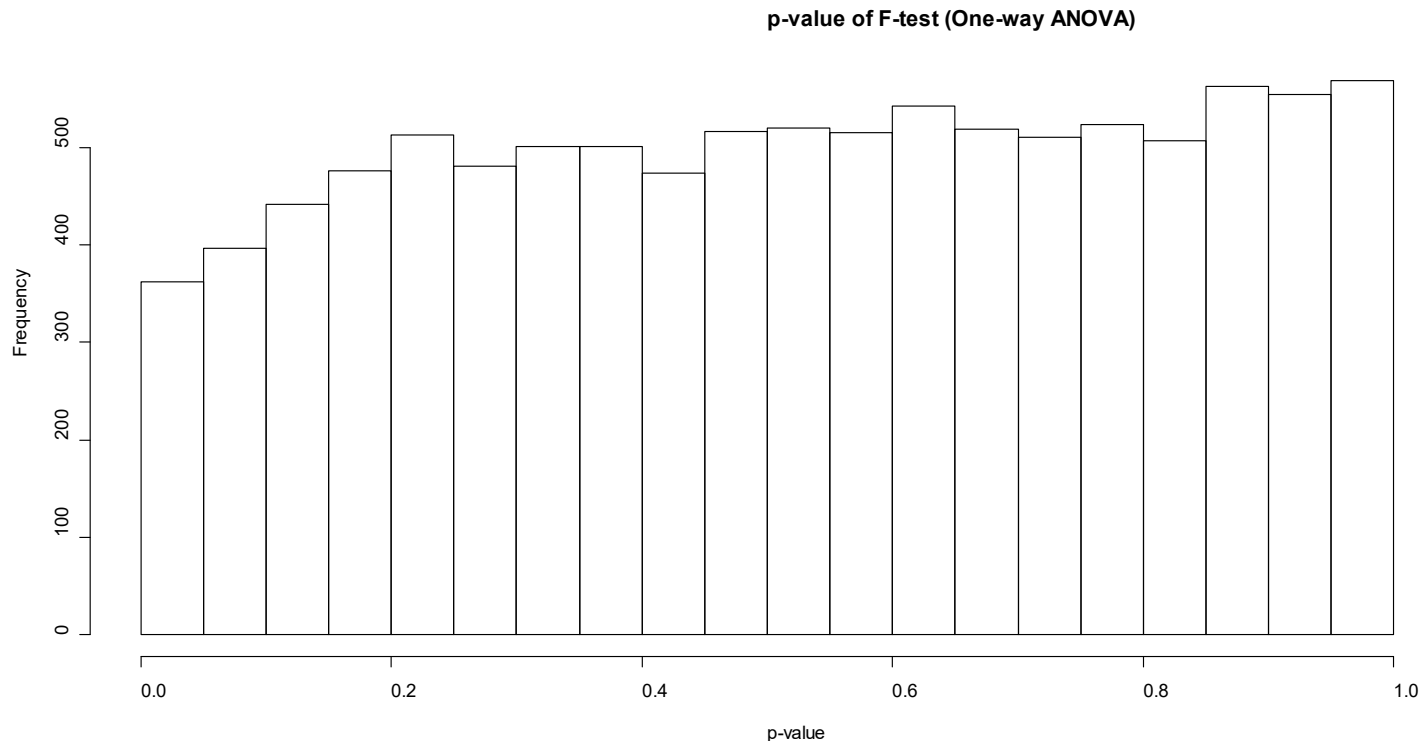


是否符合指數分配可由KS檢定確認。

範例：檢查單因子變異數分析，假設有三個因子，每組的10個觀察值皆服從變異數為一的常態分配。

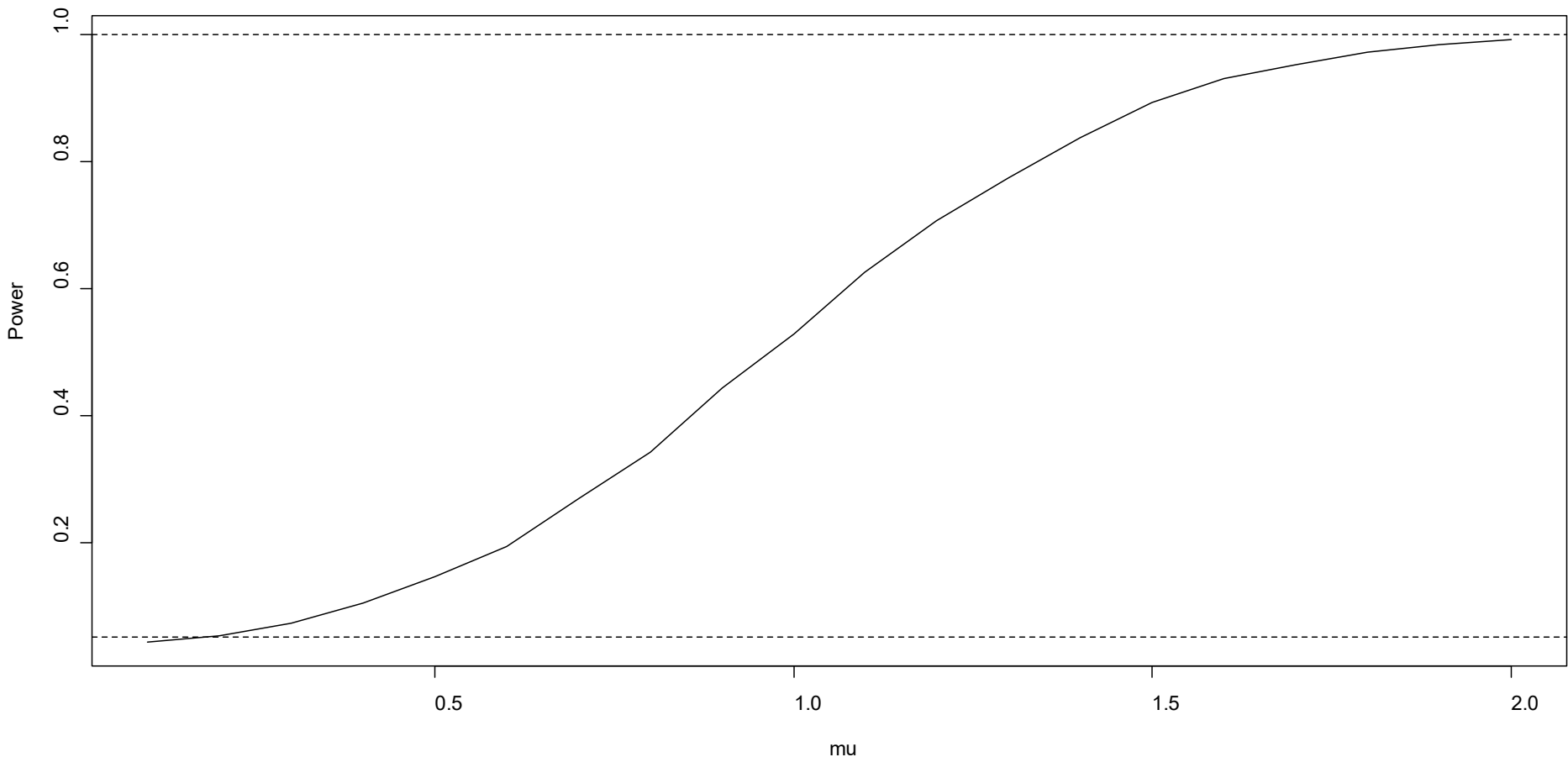
(1) 計算型一誤差：三組期望值都是0。

→ 一萬次模擬的p-value不像是均勻分配，小於0.05者約為0.0363。

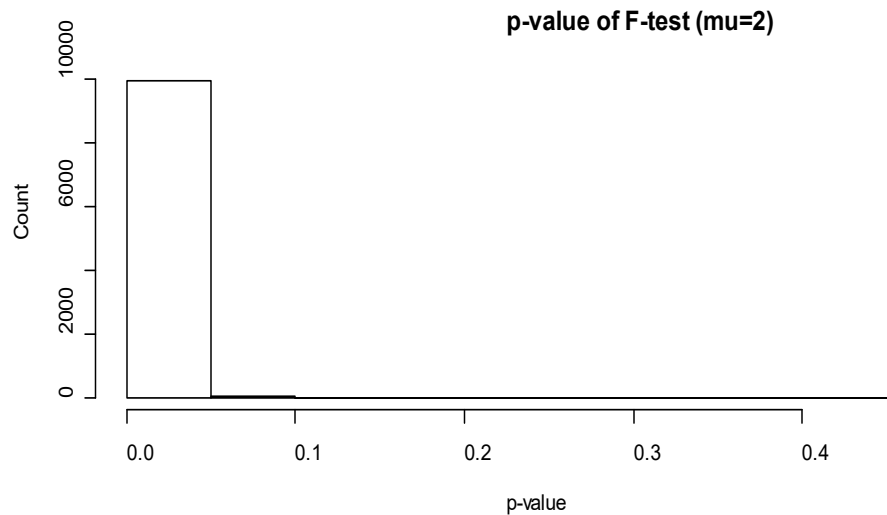
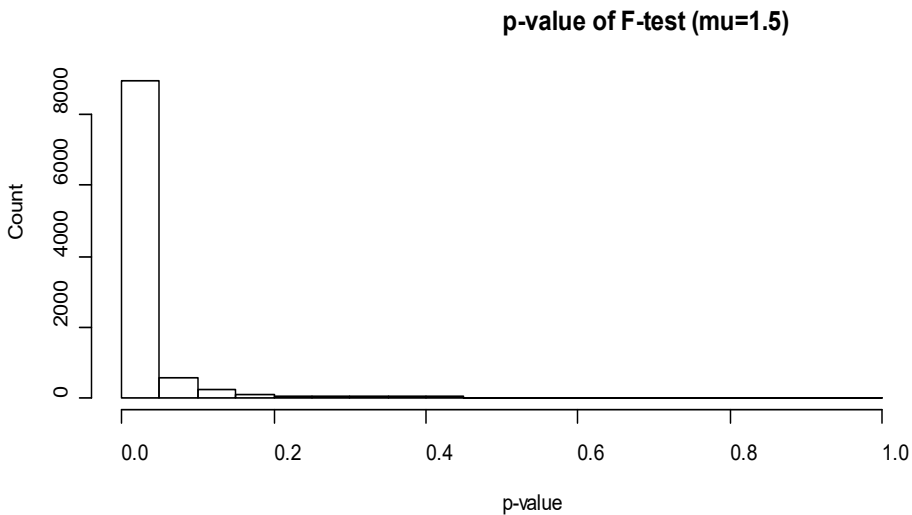
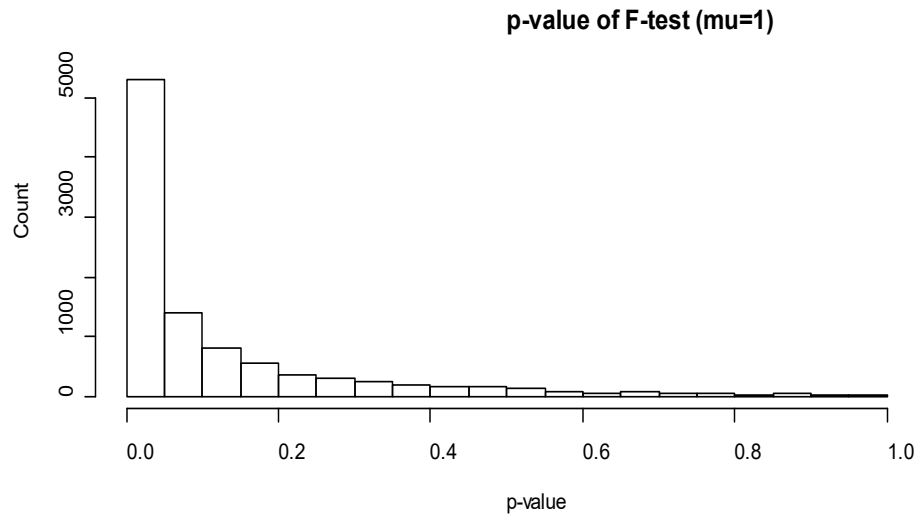
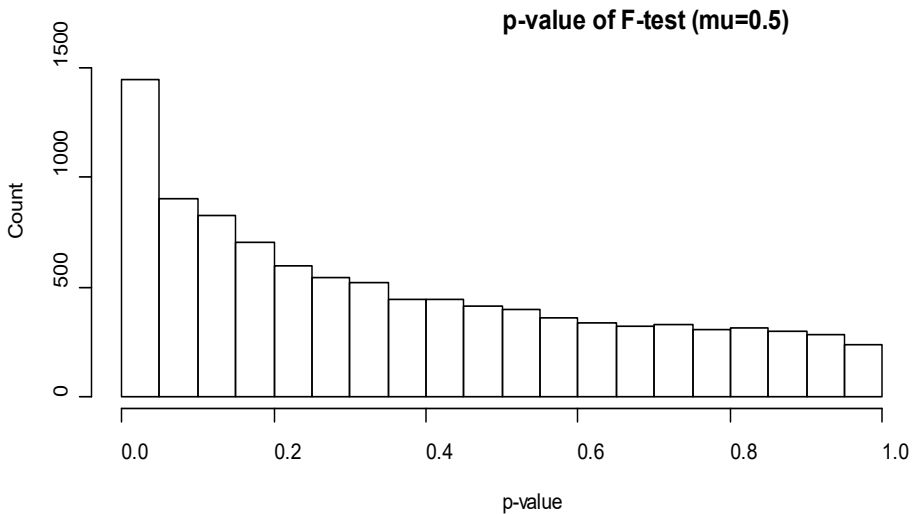


(2) 計算檢力：假設其中一組的期望值為 $\mu$ ，其他兩組為0，以下檢力曲線皆為一萬次模擬的結果：(下線為0.05！)

Power of F-test (One-way ANOVA)



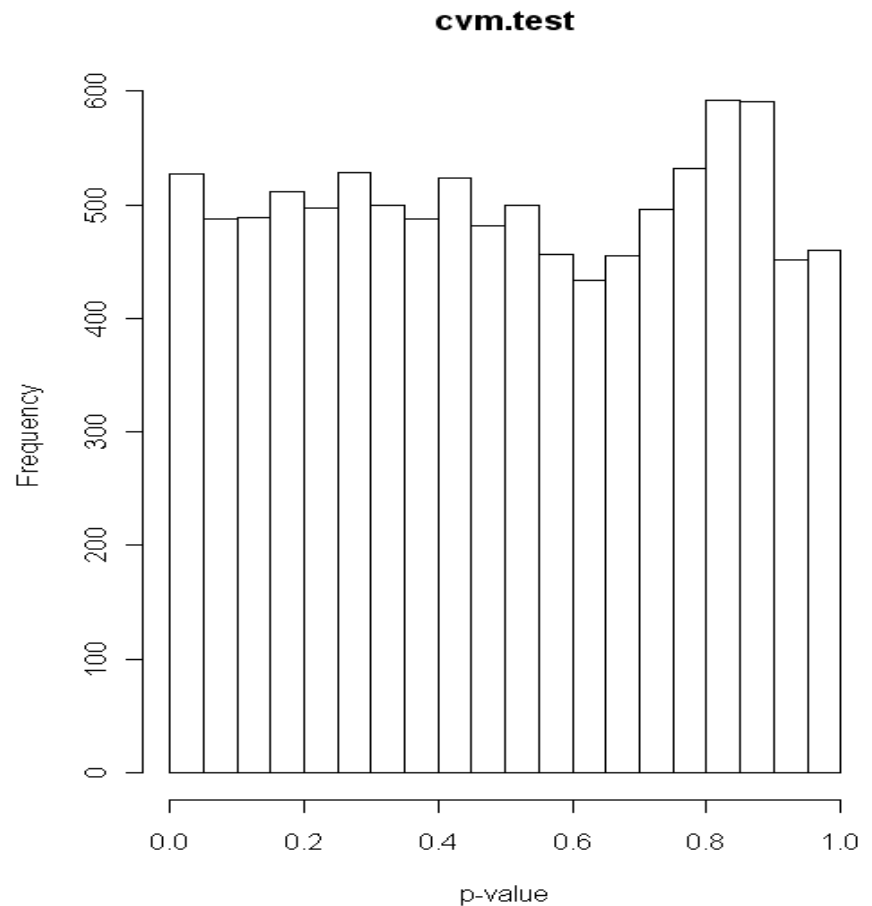
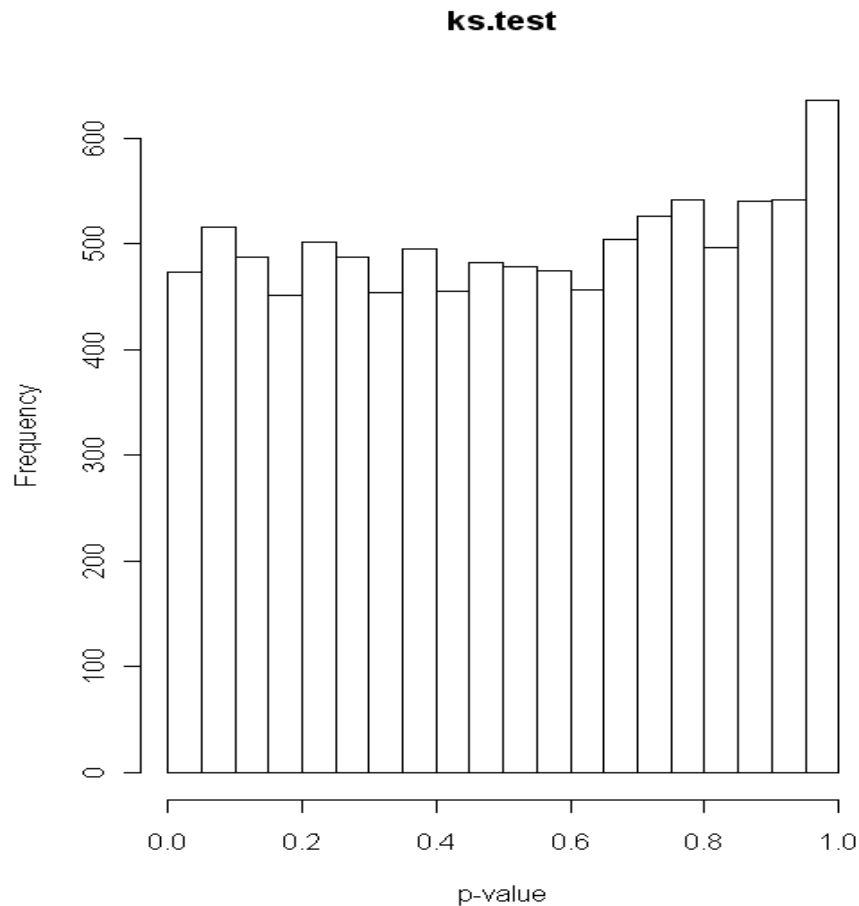
# F-test p-value在 $\mu=0.5, 1, 1.5, 2$ 的長條圖





範例：檢查R軟體的常態分配檢定，像是「ks.test」或「cvm.test」。

→ 100筆資料來自標準常態分配，計算檢定量的p-value，重複模擬一萬次。



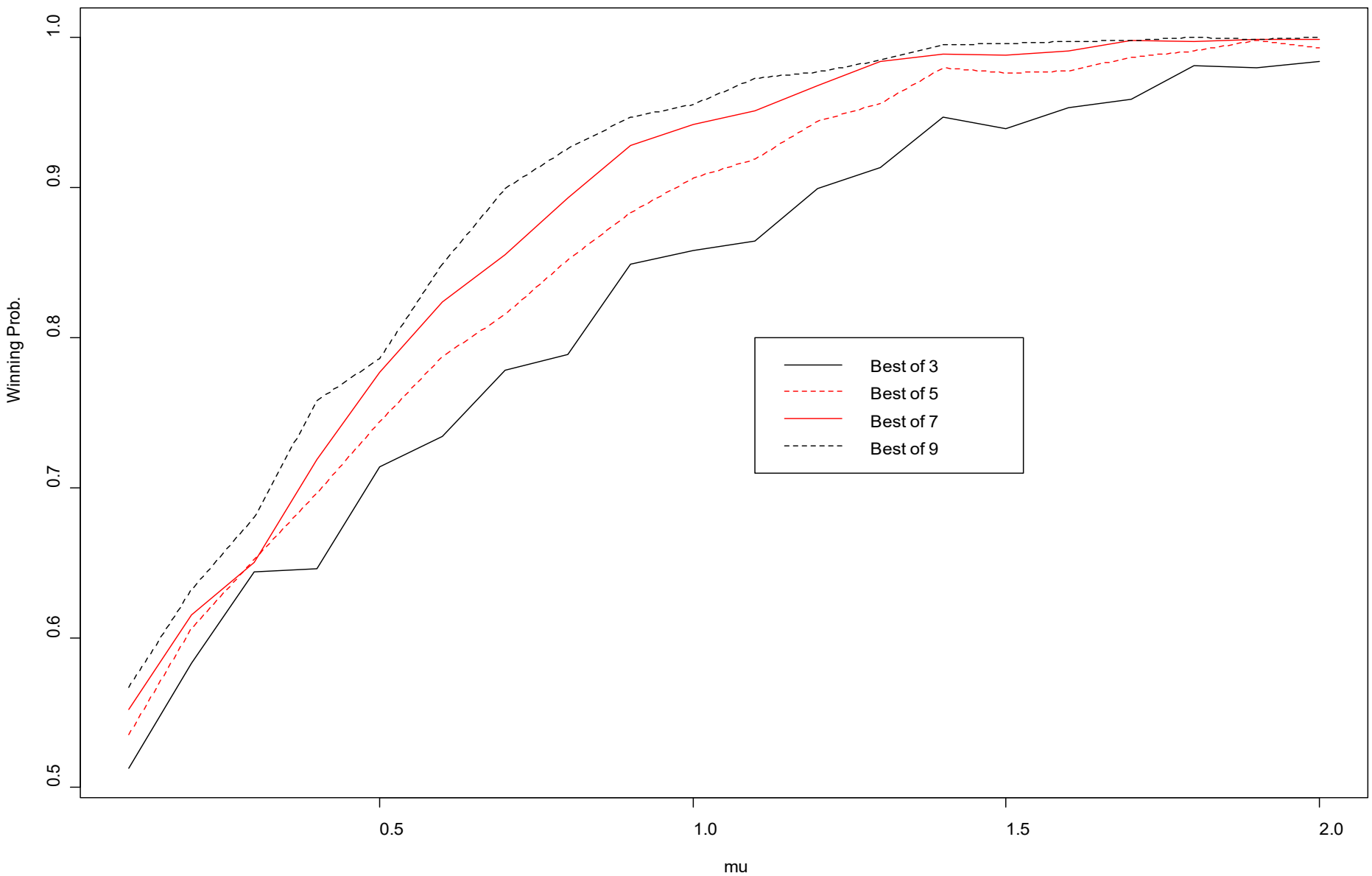
■一萬次模擬的p-value理論上應該接近均勻分配 $U(0,1)$ ，因此分別以卡方適合度檢定（分成10組）計算，發現兩者的檢定值都大於50，遠大於臨界值16.92。

→若再以ks.test檢定均勻分配，也是拒絕！

|     | 0    | 1    | 2    | 3   | 4    | 5   | 6   | 7    | 8    | 9    |
|-----|------|------|------|-----|------|-----|-----|------|------|------|
| KS  | 989  | 938  | 990  | 949 | 937  | 952 | 961 | 1068 | 1038 | 1178 |
| CVM | 1015 | 1001 | 1025 | 987 | 1005 | 956 | 888 | 1028 | 1183 | 912  |

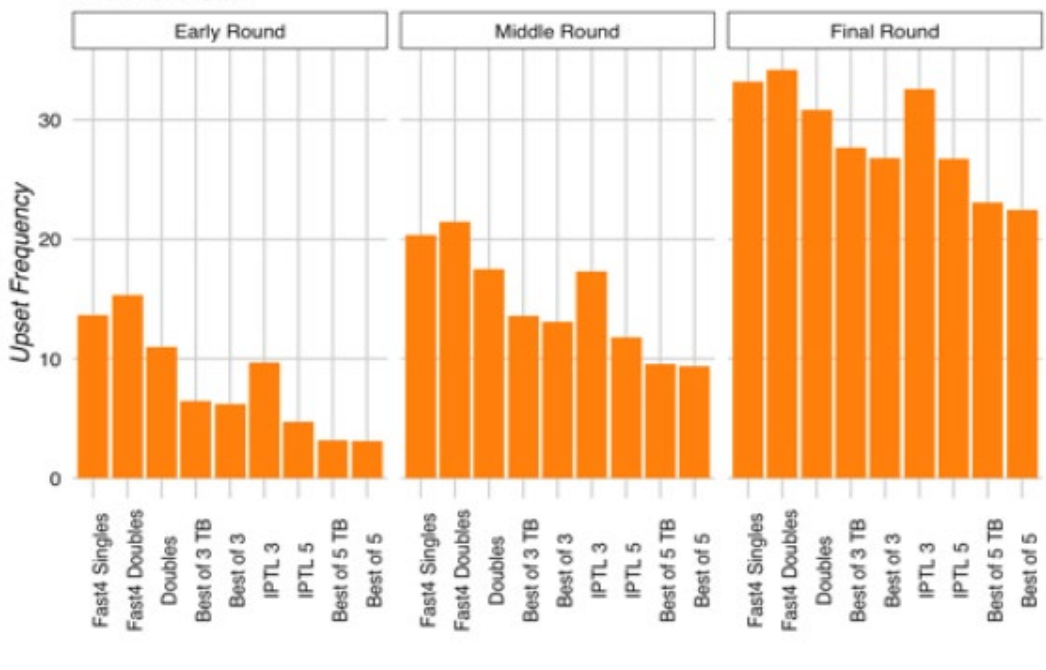
# 七戰四勝等賽制中強隊勝出的機會

# of Games vs. Winning Prob.



| Name          | Description                                                                                                                                                        |
|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Best of 3     | Sets are played to six games. Winner is first to win two sets. Tiebreak is played at 6-6 except in third set when an advantage set is played.                      |
| Best of 3 TB  | This format is the same at the Best of 3 except that third set uses a tiebreak at 6-6.                                                                             |
| Best of 5     | Same as Best of 3 but the winner is the first to win three sets.                                                                                                   |
| Best of 5 TB  | This format is the same at the Best of 5 except that third set uses a tiebreak at 6-6.                                                                             |
| Doubles       | First two sets are played as a Best of 3 but there is no ad and tiebreaks are played at 6-6. Third sets are decided by a a ten-point tiebreak.                     |
| IPTL 3        | Three sets with no ad and no lets. At 5-5, players play a first to 7-point tiebreak. <a href="http://www.iptlworld.com/format">http://www.iptlworld.com/format</a> |
| IPTL 5        | Same as IPTL but five sets.                                                                                                                                        |
| Fast4 Singles | Best of 3 sets with sets played to four games. There is no ad and no lets. At 3-3, players play a first to five-point tiebreak.                                    |
| Fast4 Doubles | Same format as Fast4 singles for first 2 sets. Third set is decided by ten-point tiebreak.                                                                         |

## Men's Tour

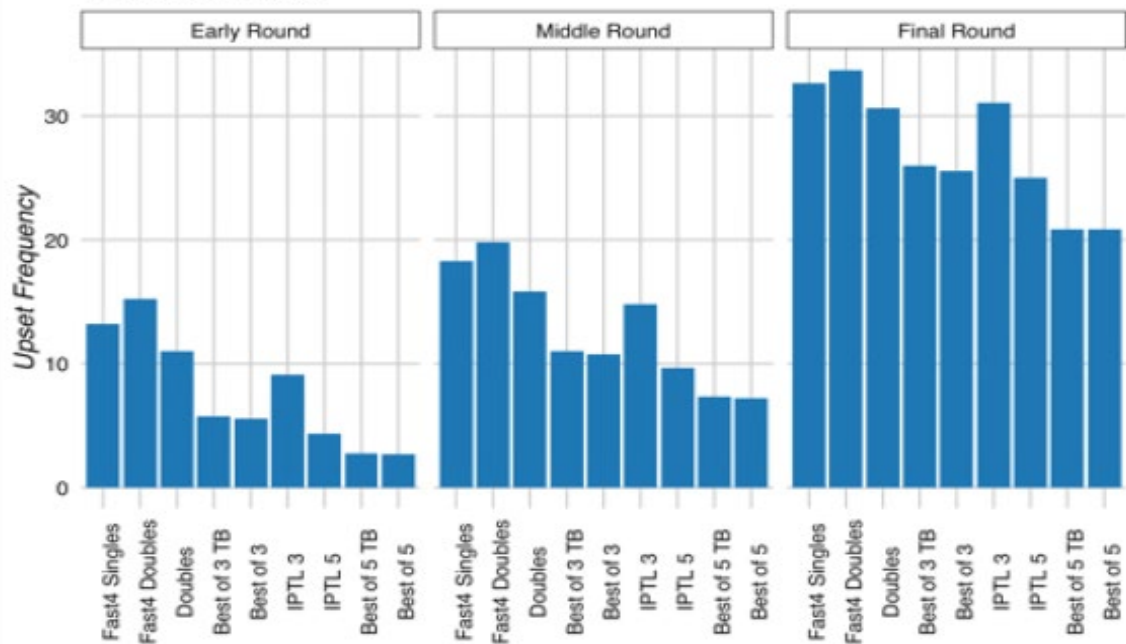


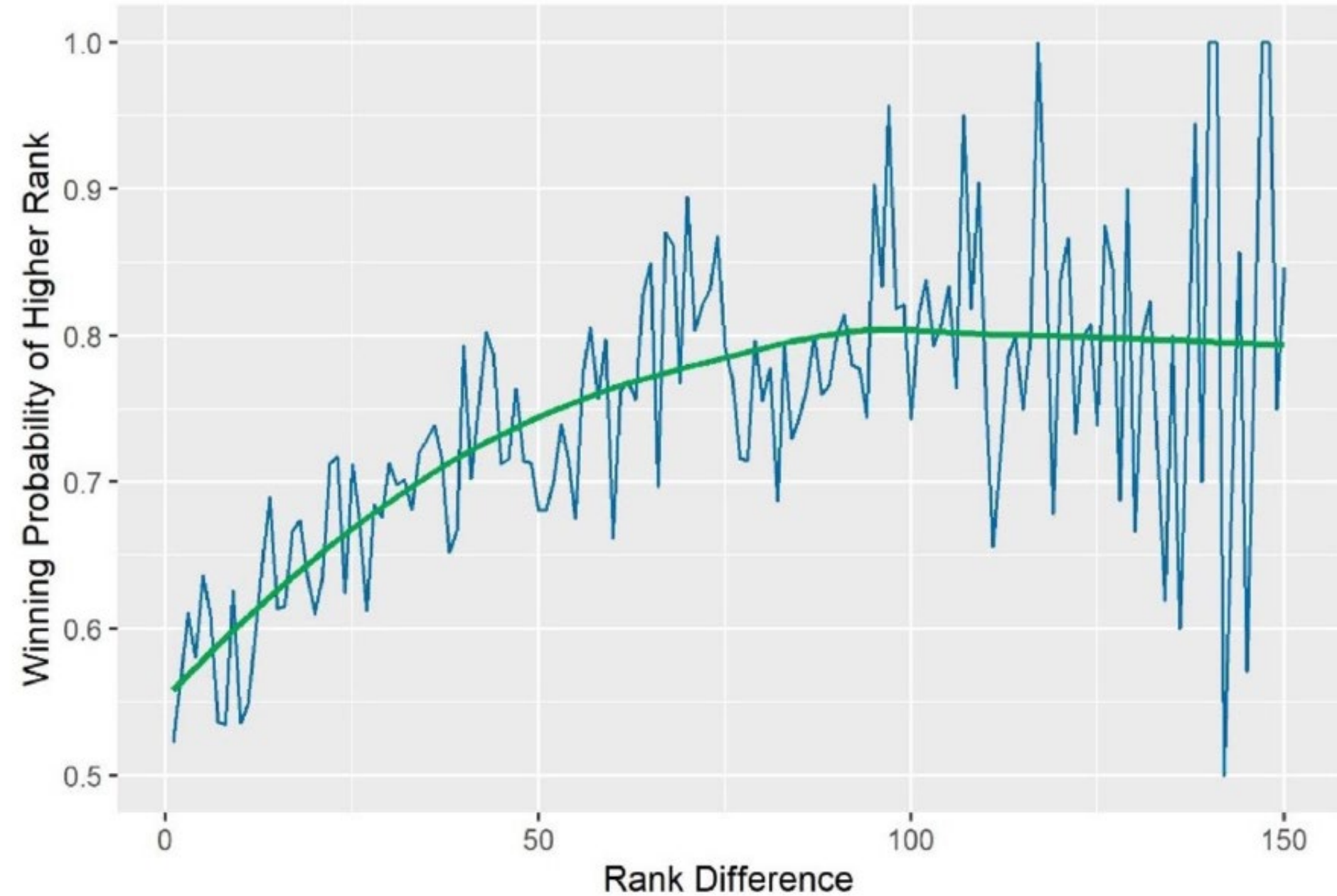
□ 盤數愈多、種子球員勝出的機會愈大。

→ 愈接近決賽、翻盤的機會也愈大。

註：實力愈接近！

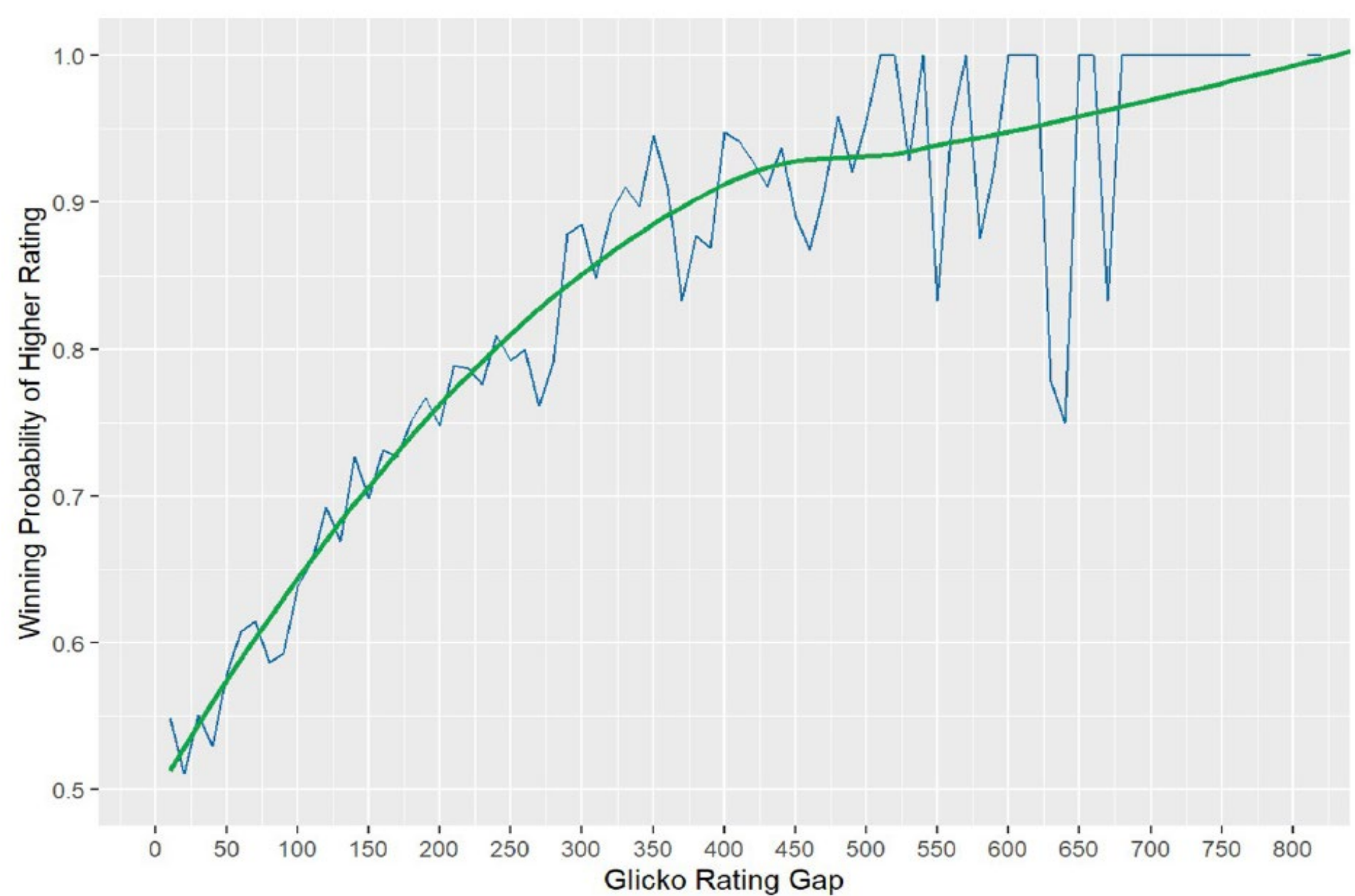
## Women's Tour





## 排名差異與勝率(四大滿貫、2000~2019)

J.C. Yue, E.P. Chou, M.H. Hsieh, and L. Hsiao (2022), "A Study of Forecasting Tennis Matches via the Glicko Model".



# Glicko分數與勝率(四大滿貫、2000~2019)

J.C. Yue, E.P. Chou, M.H. Hsieh, and L. Hsiao (2022), "A Study of Forecasting Tennis Matches via the Glicko Model".

NetLogo: Wolf Sheep Predation {/Users/tschoeh/Download/NetLogo 2.1/models/Sample Models/Biology}

Interface Information Procedures Errors

Edit Delete abc Button abc 5 Slider On Off Switch abc 5 Chooser abc 5 Monitor abc Plot abc c Output abc def ghi jkl Text

show-energy?

Grass settings

grass?
  grass-regrowth-time 30

Sheep settings

initial-number-sheep 82

sheep-gain-from-food 4.0

sheep-reproduce 4.0%

Wolf settings

initial-number-wolves 49

wolf-gain-from-food 20.0

wolf-reproduce 5.0%

time-ticks: 233
 sheep: 133
 wolves: 36
 grass / 4: 139

populations

sheep  
 wolves  
 grass / 4

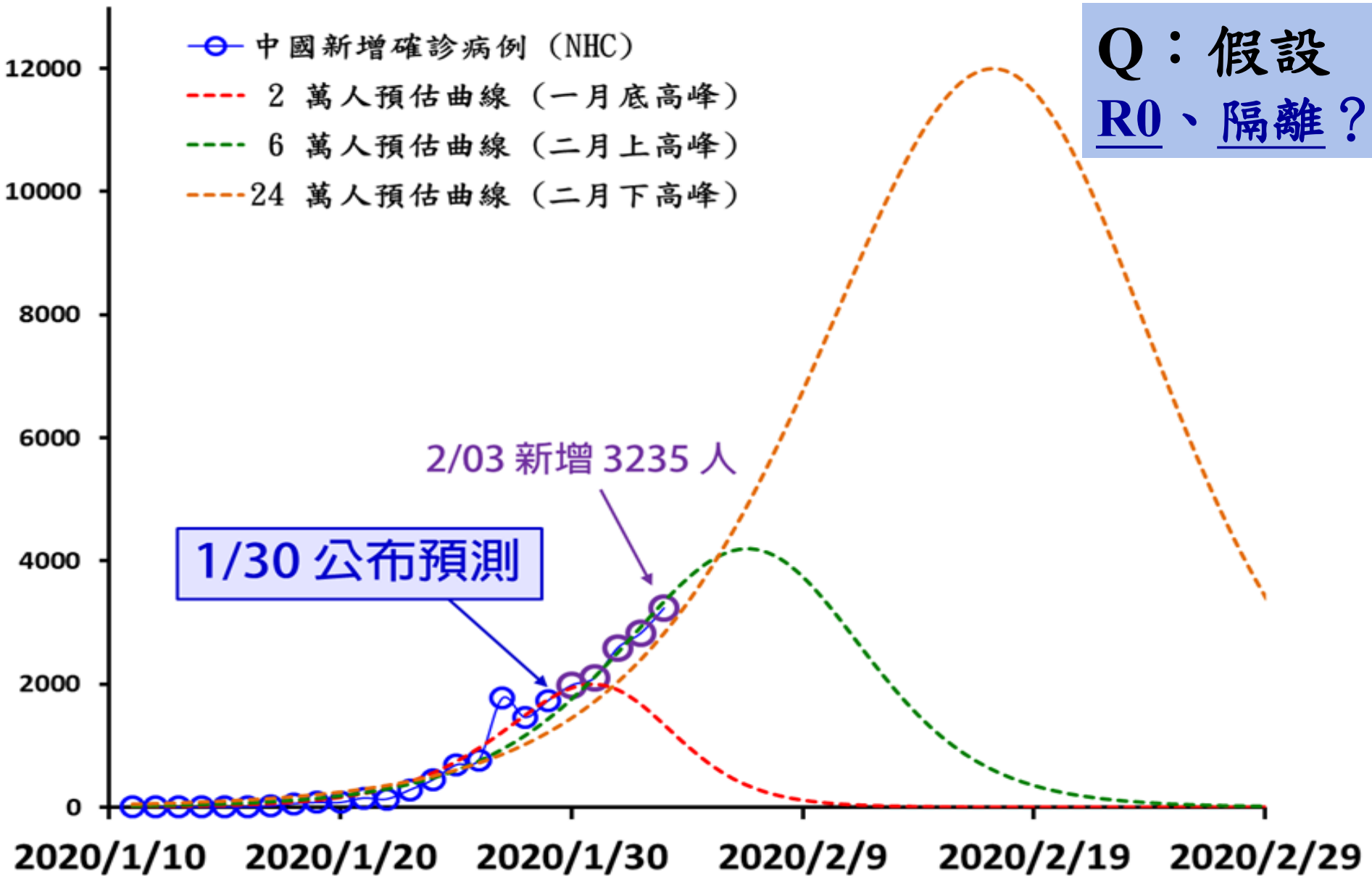
Command Center

[https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FThe-NetLogo-Predator-Prey-simulation-Wolf-Sheep-Grass\\_fig1\\_222712675&psig=AOvVaw00VcW-auVzvkgWIZ8jAScB&ust=1586753167702000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCNjKucWK4ugCFQAAAAAdAAAAABAD](https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FThe-NetLogo-Predator-Prey-simulation-Wolf-Sheep-Grass_fig1_222712675&psig=AOvVaw00VcW-auVzvkgWIZ8jAScB&ust=1586753167702000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCNjKucWK4ugCFQAAAAAdAAAAABAD)

狼、羊、草地的自然均衡 (Predator & Prey 模擬)



# COVID-19 案例最高峰 (台大化學系徐承志)



Q: 假設  
R0、隔離?

# 疾病的基本傳染數(Basic Reproduction Number, R0)

| 疾病                    | 病死率    | 傳染途徑  | R0        |
|-----------------------|--------|-------|-----------|
| 麻疹                    | 1-3%   | 空氣傳播  | 12-18     |
| 天花                    | 95%    | 空氣、飛沫 | 5-7       |
| 愛滋病<br>(不經治療)         | 80-90% | 性行為傳播 | 2-5       |
| 霍亂                    | 1-50%  | 糞口傳播  | 1.06-2.63 |
| 伊波拉(2014)             | 83-90% | 體液傳播  | 1.5-2.5   |
| 普通流感                  | 0.5%   | 空氣、飛沫 | 0.9-2.1   |
| 流行性感冒<br>(1918 流感大流行) | >2.5%  | 空氣、飛沫 | 2-3       |
| SARS (2003)           | 11%    | 空氣、飛沫 | 0.85-3    |

COVID-19

2~3%

飛沫、接觸

1.4-3.8

# 武漢大逃亡 Worldpop 分析路線圖

