統計計算與模擬

政治大學統計系余清祥

2023年2月21日

第一單元：模擬

http://csyue.nccu.edu.tw

什麼是模擬？

- 模擬(Simulation)：模仿虛擬

→Feign, pretend to have or feel; pretend to be, act like, resemble, …(《大英百科全書》)

→The practice of mimicking some or all of the behavior of one system with a different, dissimilar system.

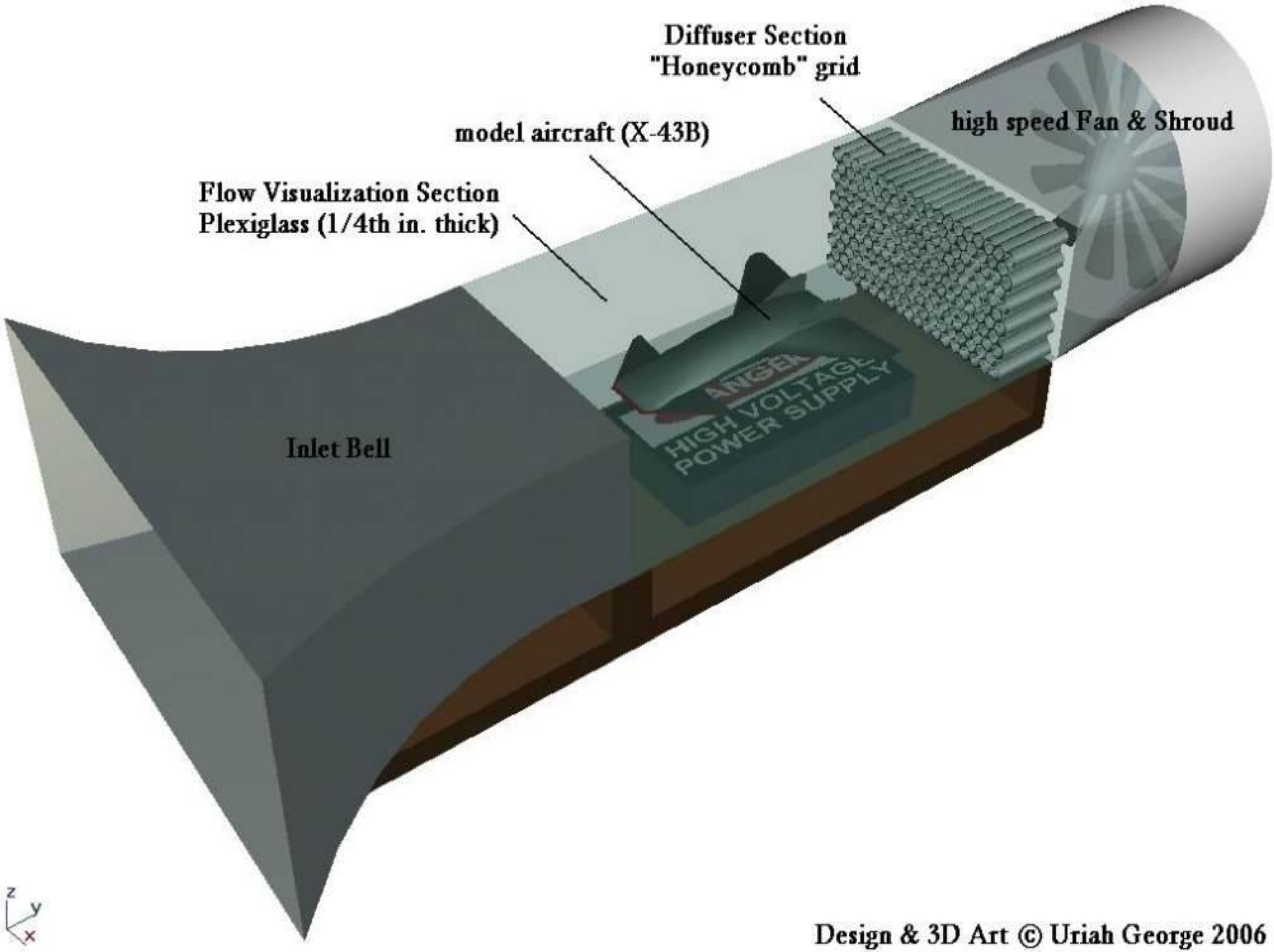
- 模擬與我們日常生活的關係

→風洞(Wind-tunnels)：汽車與飛機

→飛行模擬器(Flight simulator)

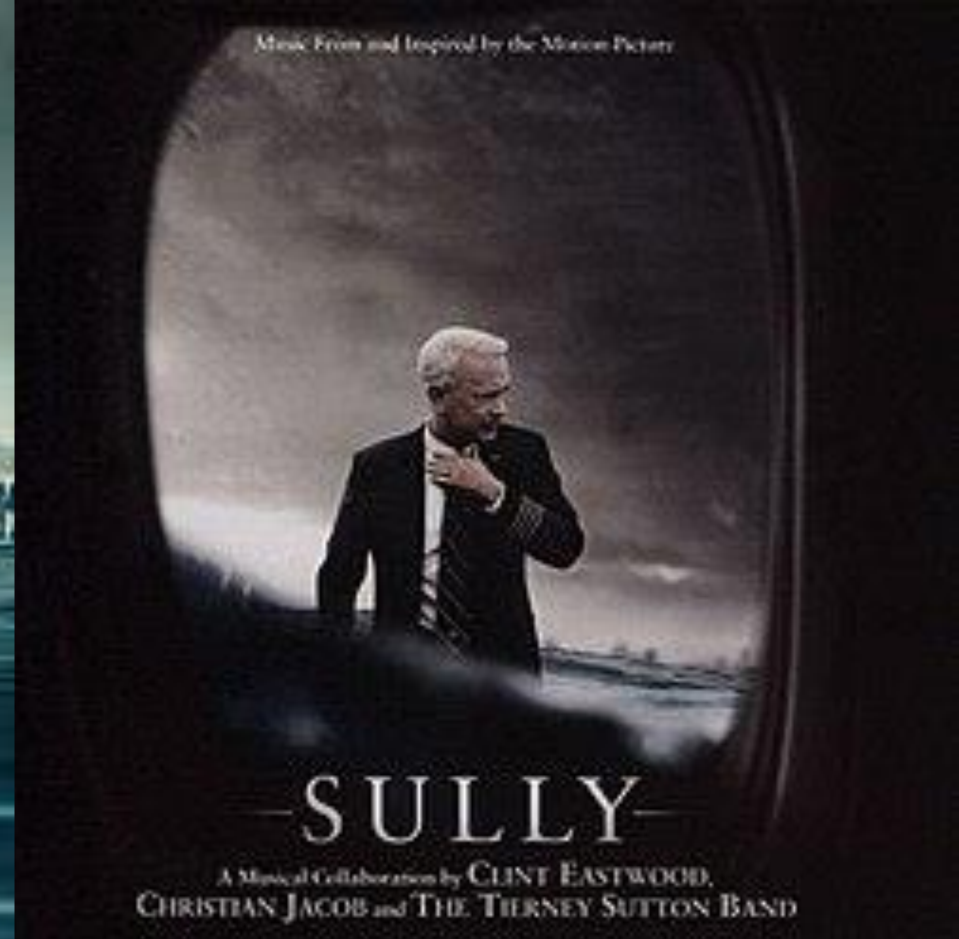
→第一個生命的誕生(閃電高熱及電擊)

→ CPU的研發(*Simulating a chip design through software programs that use models to replicate how a device will perform in terms of timing and results.*)



Diffuser Section
"Honeycomb" grid
high speed Fan & Shroud
model aircraft (X-43B)
Flow Visualization Section
Plexiglass (1/4th in. thick)
HIGH VOLTAGE
POWER SUPPLY
Inlet Bell
Design & 3D Art © Uriah George 2006

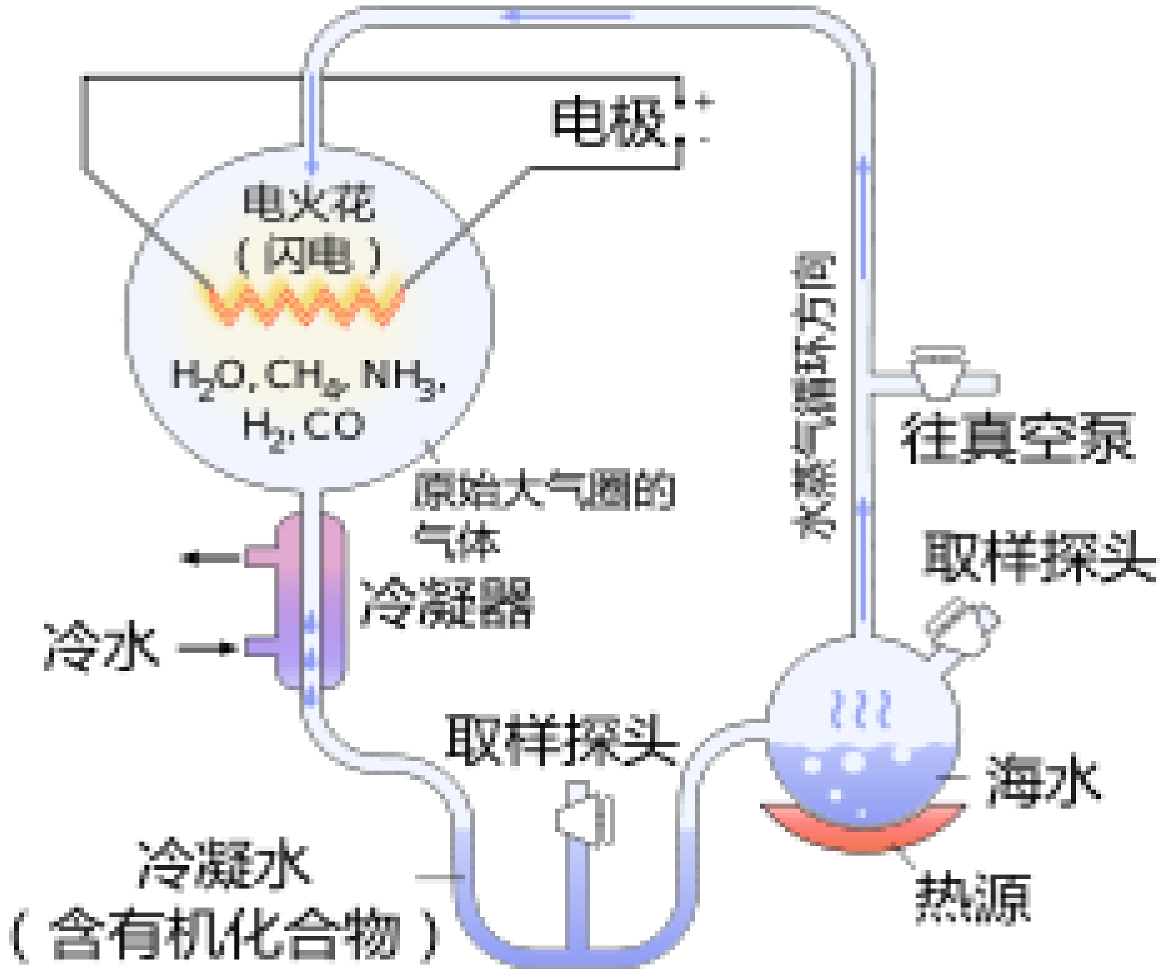

ELEC
ELEC
BAT VOLTAGE
CAB PRESSURE
AIR
AIR
FLAPS



《薩利機長：哈德遜奇蹟》

(Sully: Miracle on the Hudson)

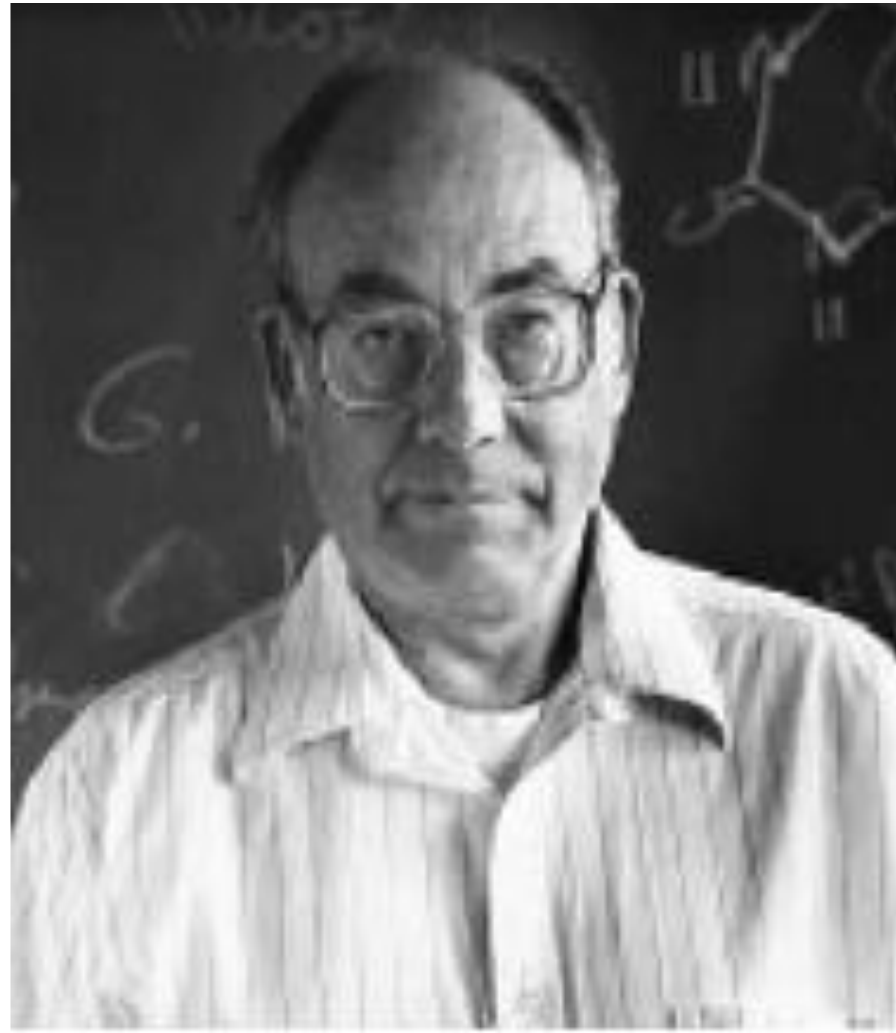
米勒-尤里實驗（Miller-Urey experiment）



註：右下燒瓶模擬海洋環境，左上燒瓶則模擬閃電。（維基百科）

米勒與尤里的生命起源實驗

- 1953年芝加哥大學米勒(Stanley Lloyd Miller, 1930-2007)在老師尤里(Harold Clayton Urey, 1893-1981)的實驗室。



圖片來源：http://www.daviddarling.info/encyclopedia/M/Miller.html

- 模擬廣為產官學各領域專家使用，已是解決問題的常見工具。

→商業的風險分析(Risk analysis)

→經濟學的政策模擬(Policy simulation)

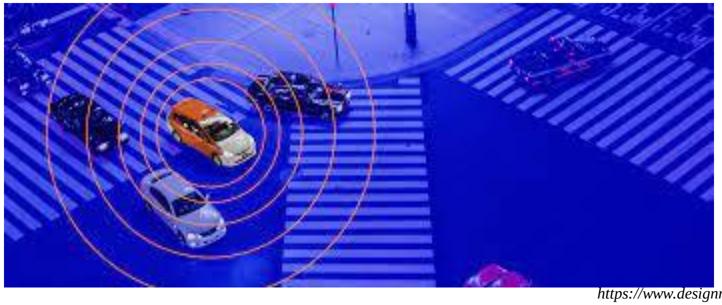
→都市計畫的交通流量(Traffic flow)、氣象預測(Weather forecasting)

→積體電路設計(CPU simulation)

→商品測試(Robot and Model car simulators)

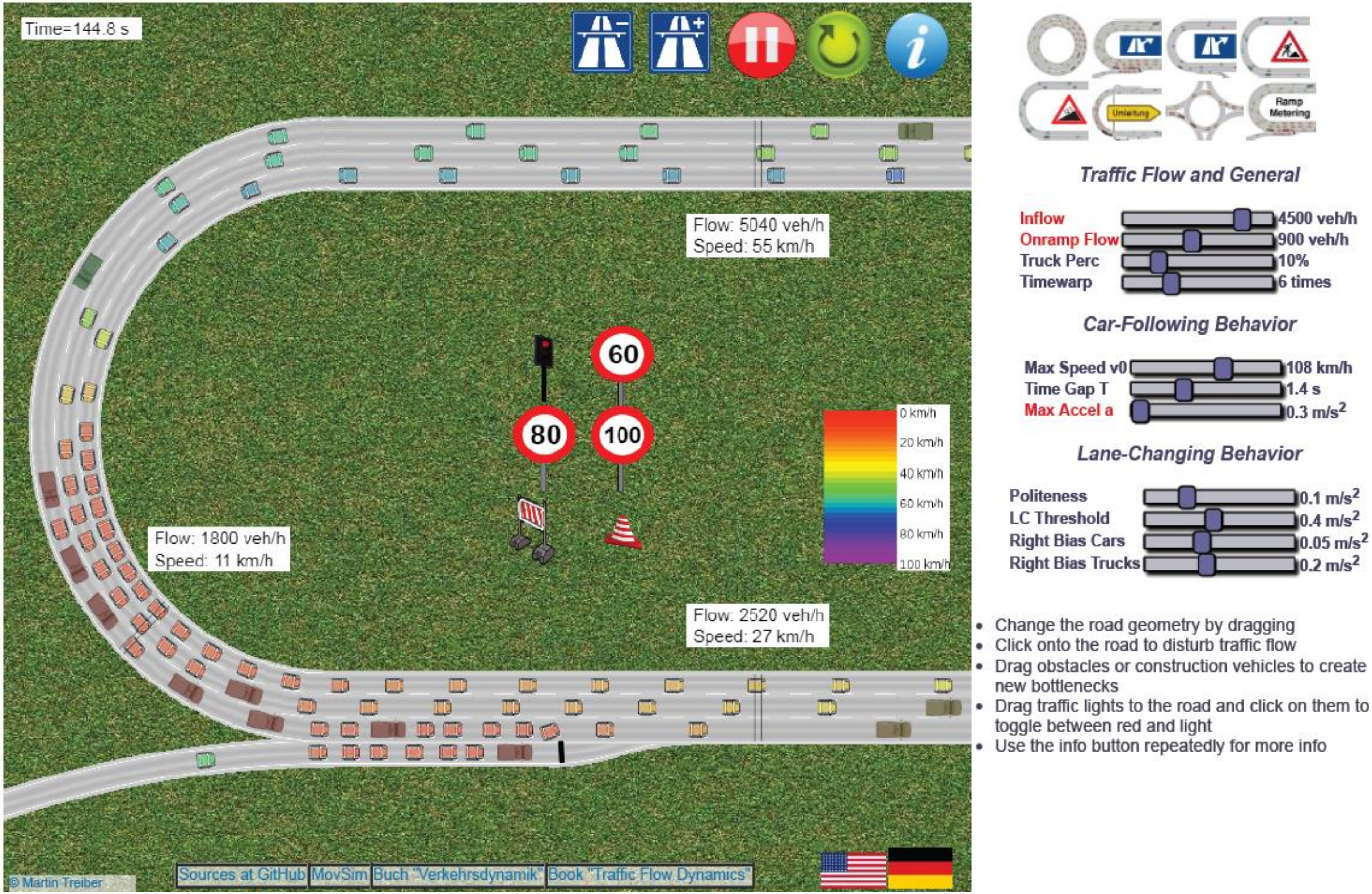
→癌症治療(Caner treatment)

https://sitn.hms.harvard.edu/flash/2017/self-driving-cars-technology-risks-possibilities/



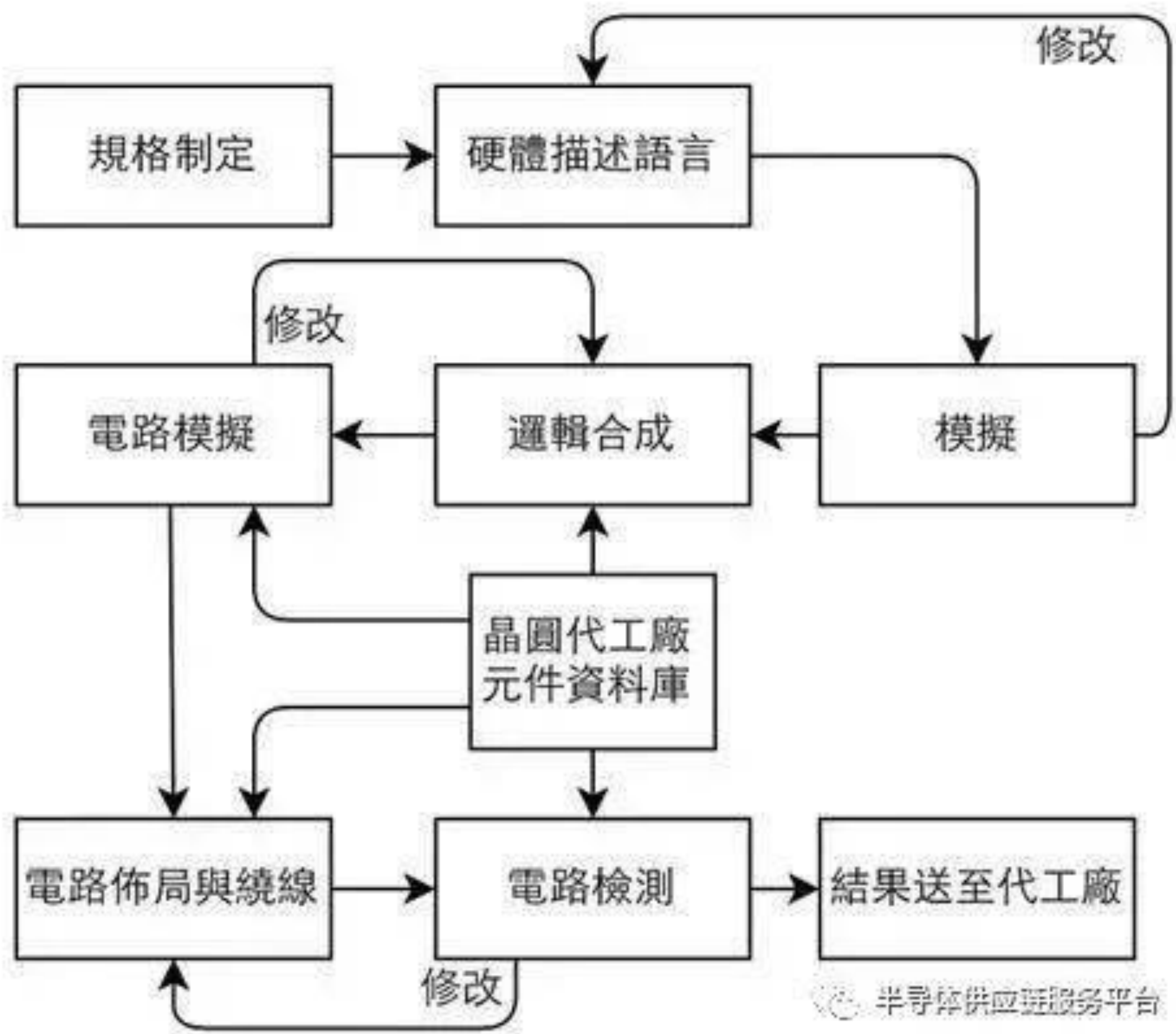
https://www.designnews.com/sites/designnews.com/files/styles/article_featured_standard/public/Design%20News/00-autonomous_SIEMENS.jpg?itok=gew1P7q7

德國的交通流量模擬

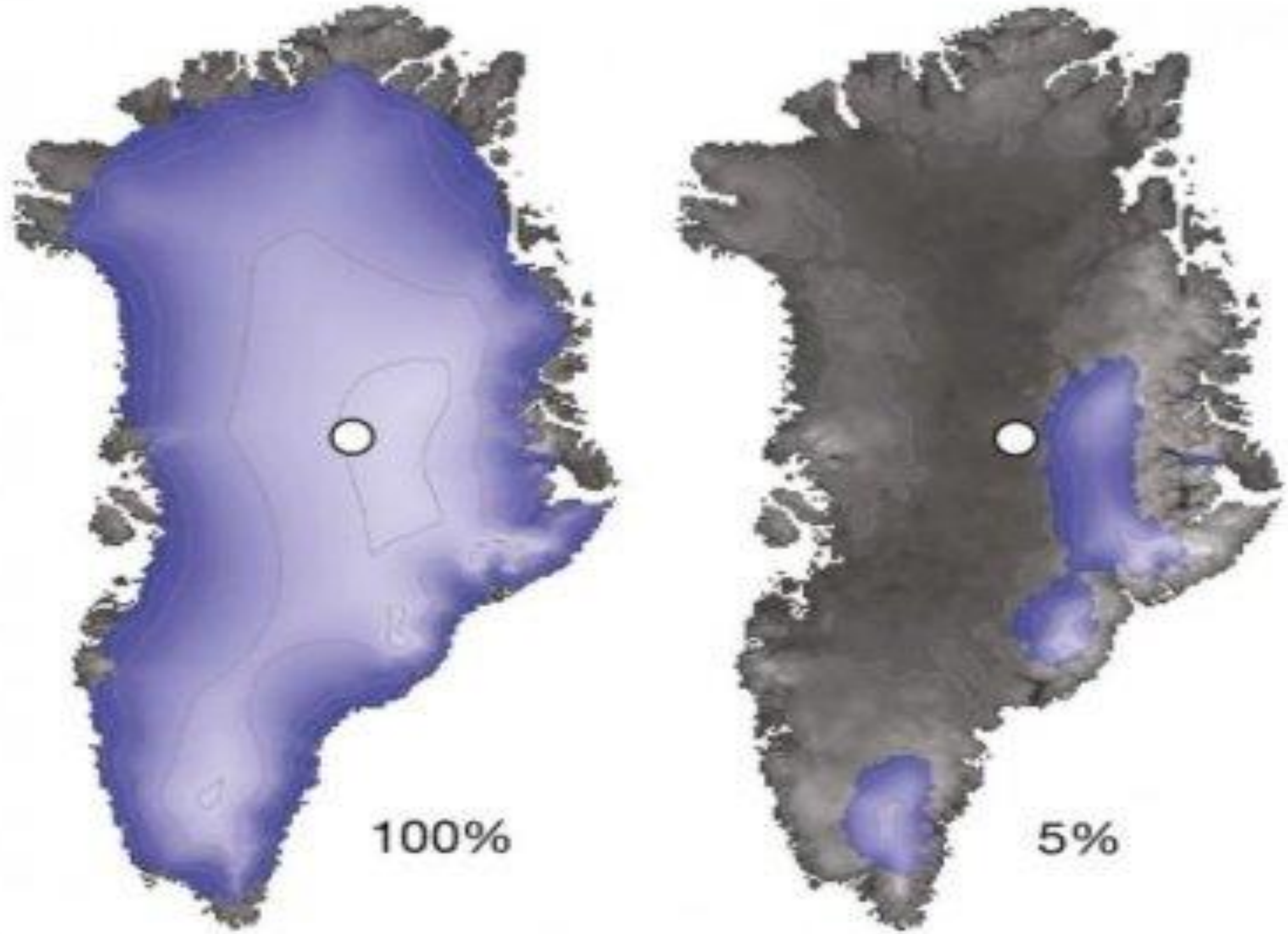


https://www.traffic-simulation.de/

工程師設計電腦芯片的流程



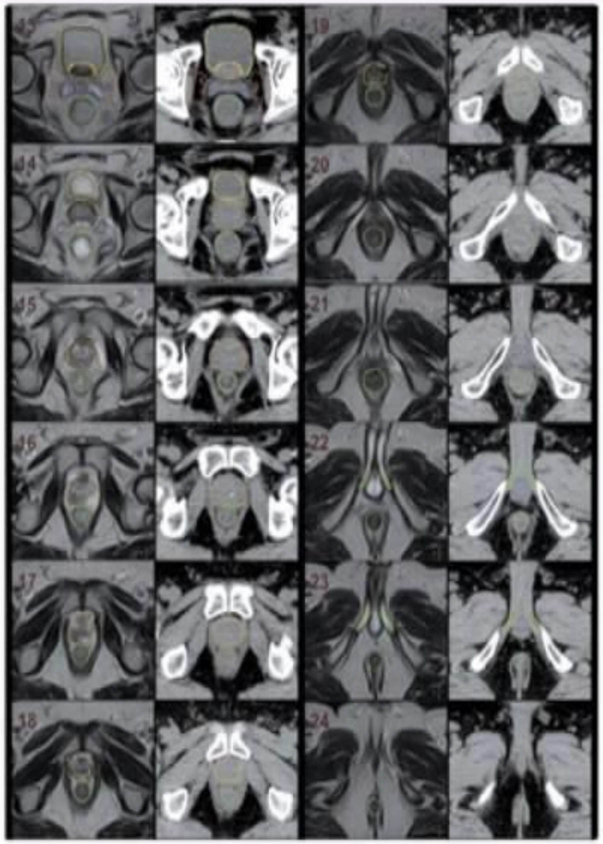
https://www.eecourse.com/article/411



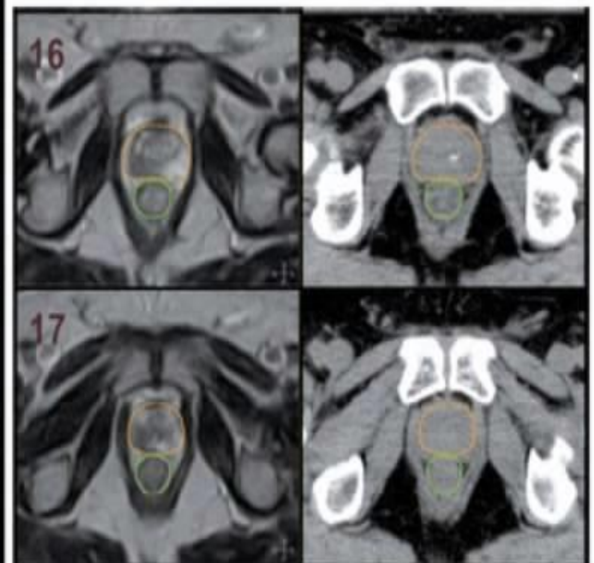
Scientists drilled nearly two miles down through the summit of the Greenland ice sheet (white dot, left), to reach bedrock. Isotopes found in the rock indicate that this site and most of Greenland were nearly ice free (right) during the recent geologic past.
(*Credit: Schaefer et al., Nature, 2016)*
https://www.sciencedaily.com/releases/2016/12/161207133453.htm

Simulation or Mapping Procedure

CT-MRI Simulation



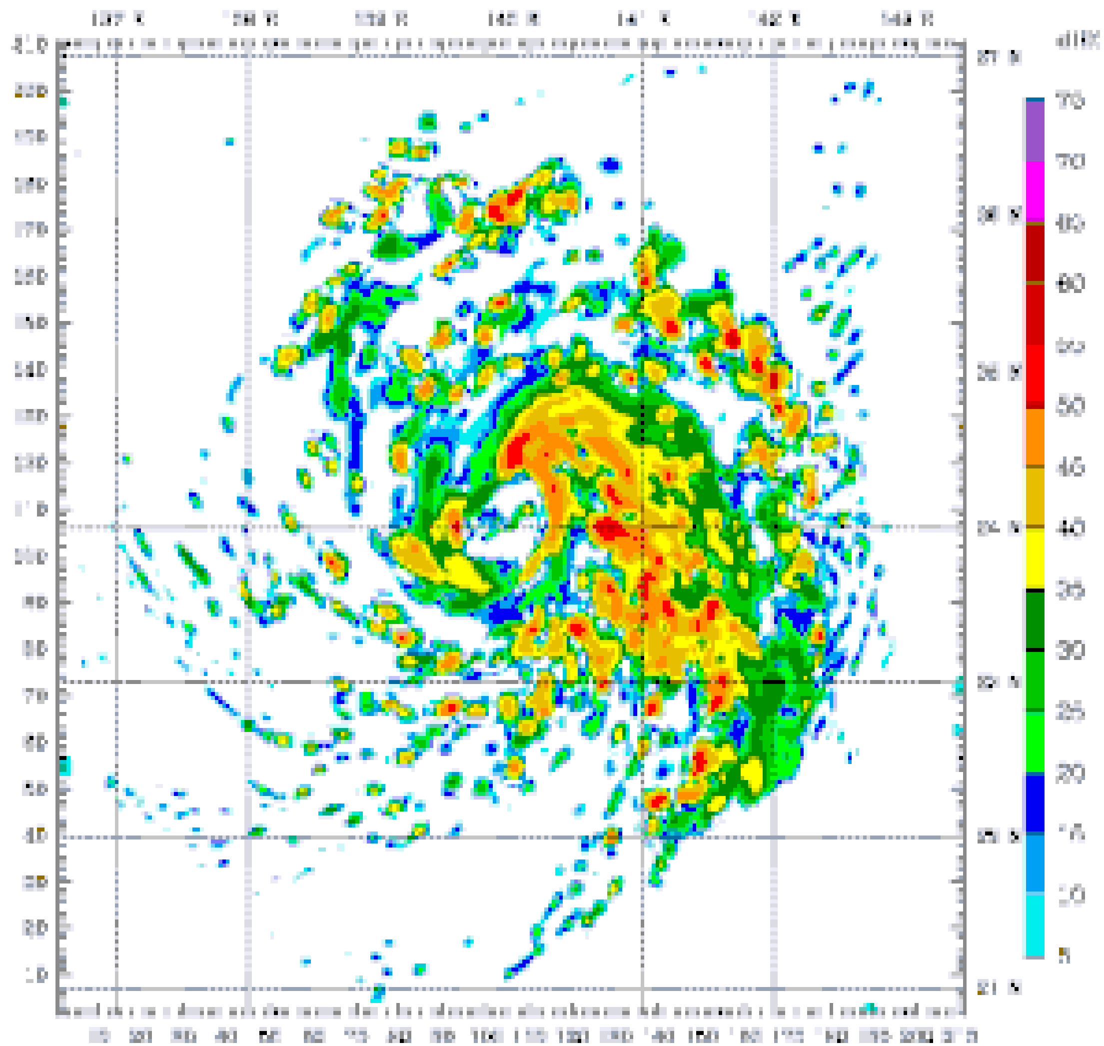
CT-MRI Fusion



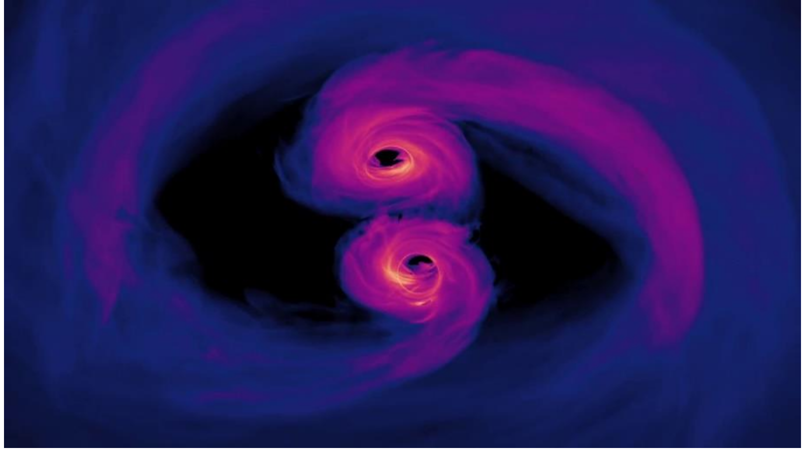
Contouring

Source: Stanford Radiation Oncologist Explains Radiation Therapy Treatment for Prostate Cancer

A 48-hour computer simulation of Typhoon Mawar

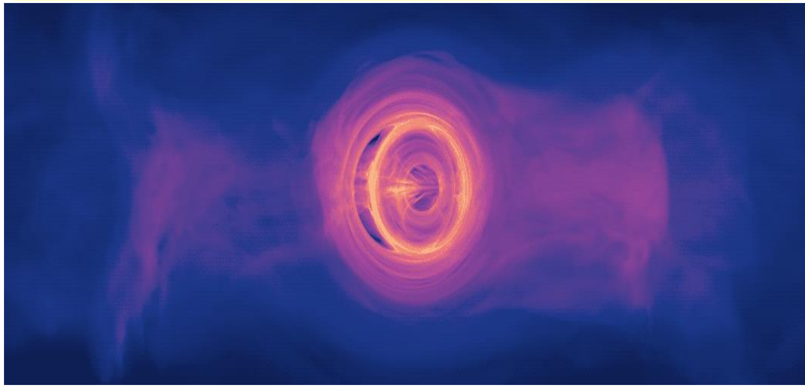


https://upload.wikimedia.org/wikipedia/commons/8/8b/Typhoon_Mawar_2005_computer_simulation_thumbnail.gif



https://i.ytimg.com/vi/i2u-7LMhwvE/maxresdefault.jpg

NASA Simulates Two Supermassive Black Holes Spiraling Toward a Collision



http://www.nasa.gov/sites/default/files/thumbnails/image/smbhb_rotate_banner.gif

蒙地卡羅法(Monte Carlo Methods)

- 電腦模擬方法多以蒙地卡羅法稱之，意謂模擬的資料與統計方法的使用。

→ 資料的模擬牽涉亂數的產生(Random Number Generation)。

- 何謂亂數？
- 為什麼稱做蒙地卡羅法？

- 為什麼要使用電腦模擬？

→ 成本較低

→ 操作較為容易

→ 無法進行真正的實驗

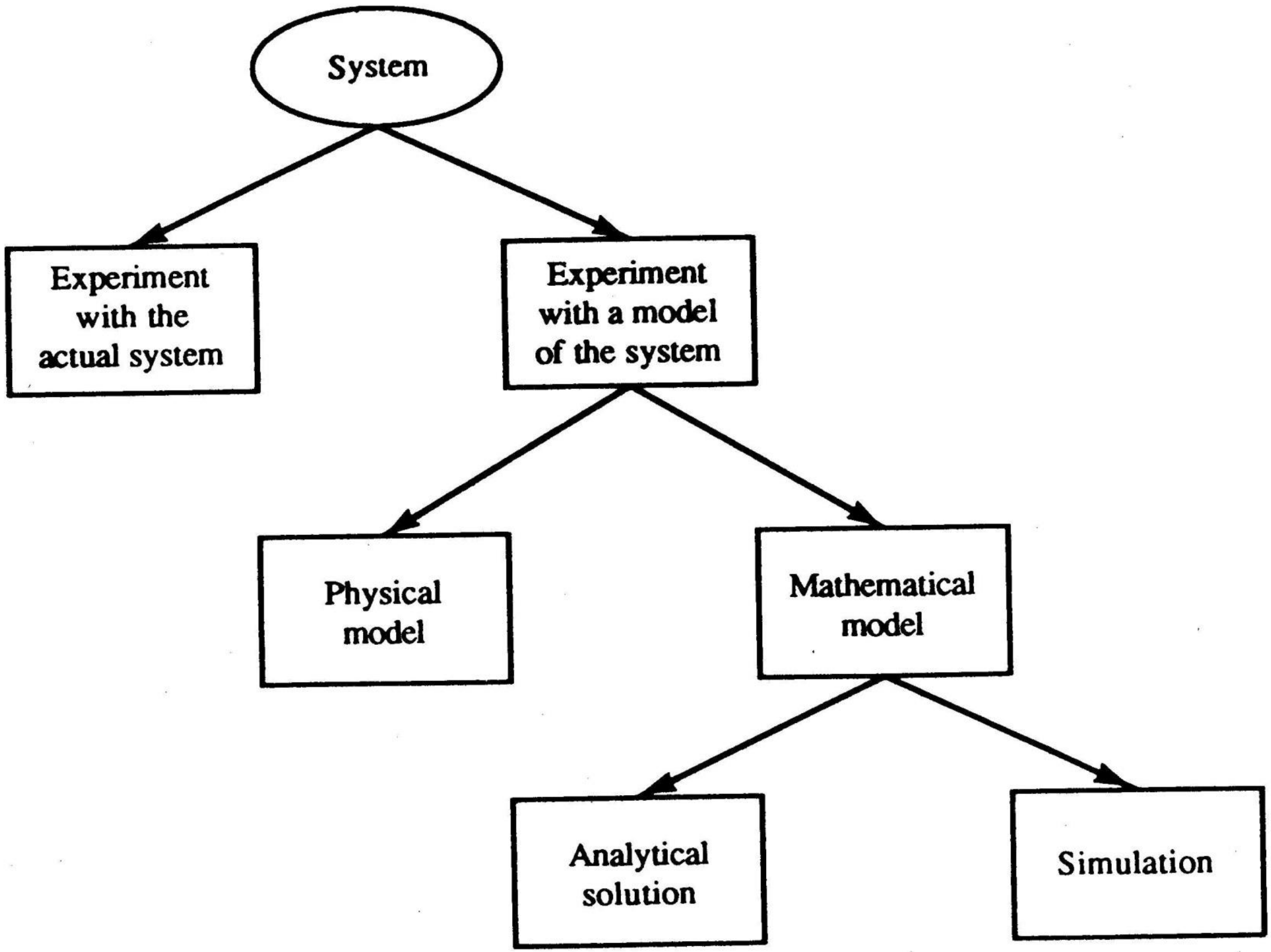
- 統計分析的三大利器

→ 數理統計(Mathematical Statistics)

→ 實驗設計(Experimental Design)

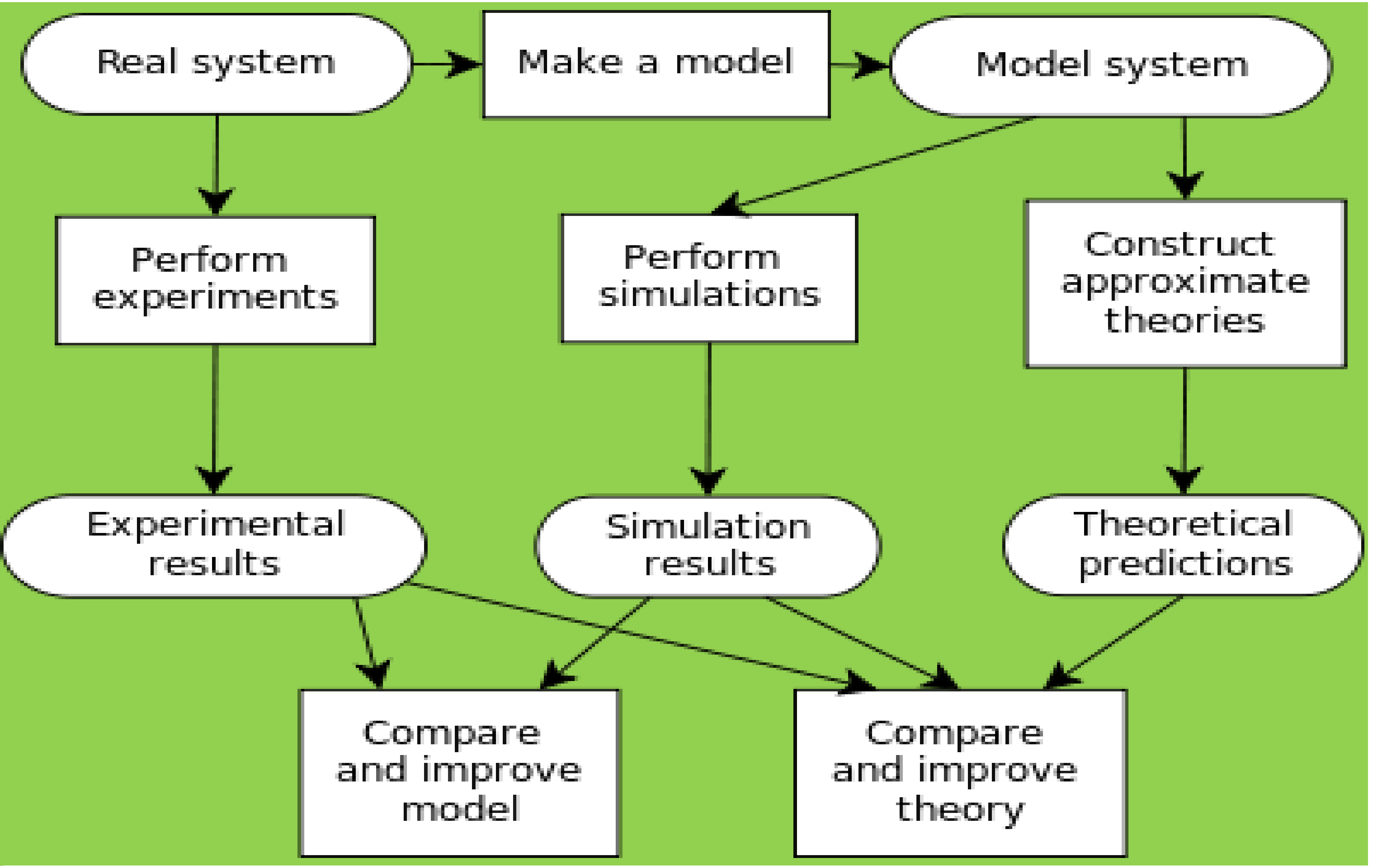
→ 電腦模擬(Computer Simulation)

系統與模擬



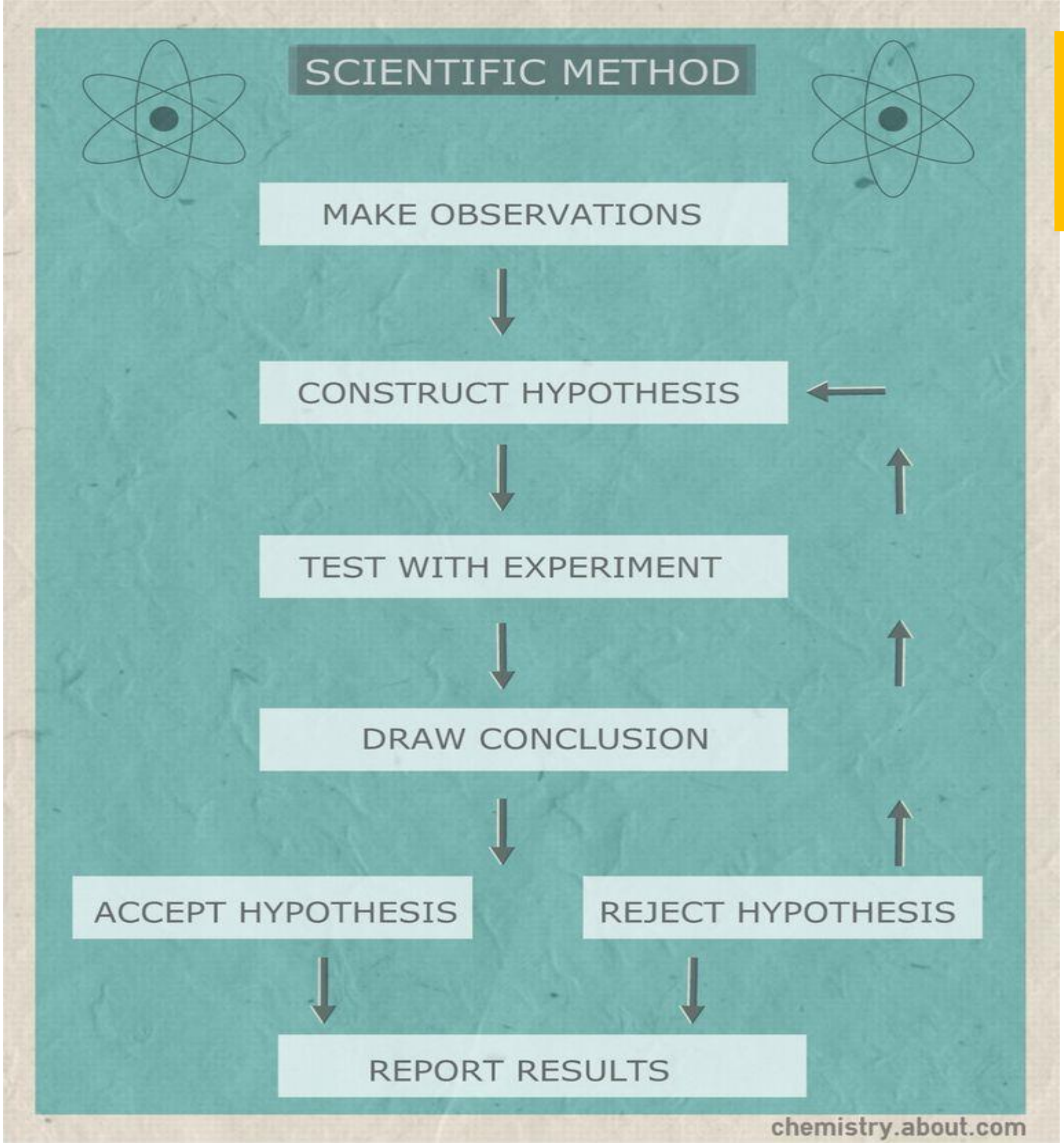
System
Experiment with the actual system
Experiment with a model of the system
Physical model
Mathematical model
Analytical solution
Simulation

系統與模擬(另一種詮釋)



https://upload.wikimedia.org/wikipedia/commons/thumb/9/9b/Molecular_simulation_process.svg/600px-Molecular_simulation_process.svg.png

科學方法
示意圖



https://ia.eferrit.com/ia/145c2ec6b919342e.jpg

電腦模擬需謹慎進行！

- Three steps should be used to produce accurate simulation models: **calibration**, **verification**, and **validation**.

→A base model should be created and calibrated so that it matches the area being studied. The calibrated model should then be verified to ensure that the model is operating as expected based on the inputs.

—摘錄自「維基百科」

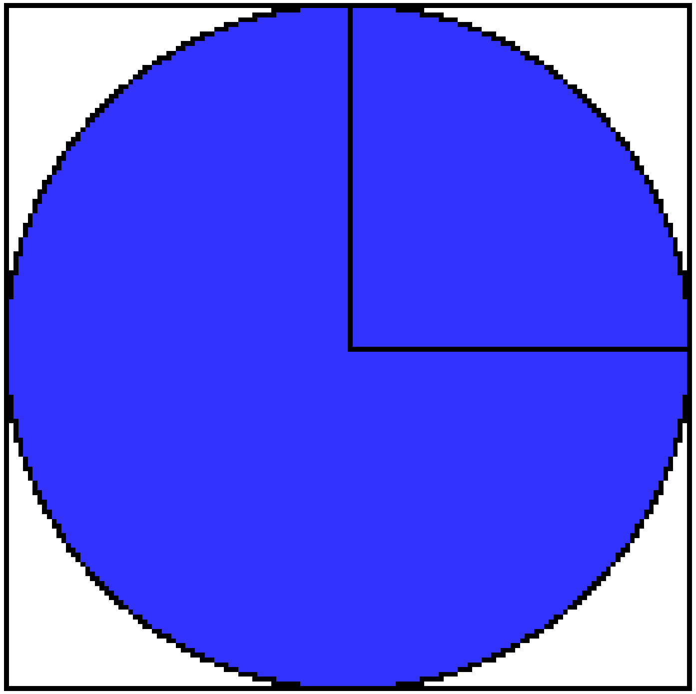
模擬分析的要件

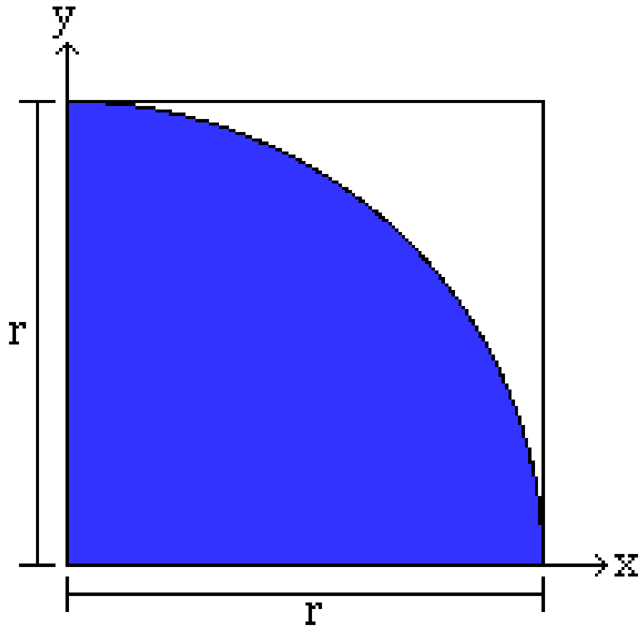
- Probability distribution function
- Random number generator
- Sampling rule
- Scoring/Tallying
- Error estimation
- Variance Reduction techniques
- Parallelization/Vectorization

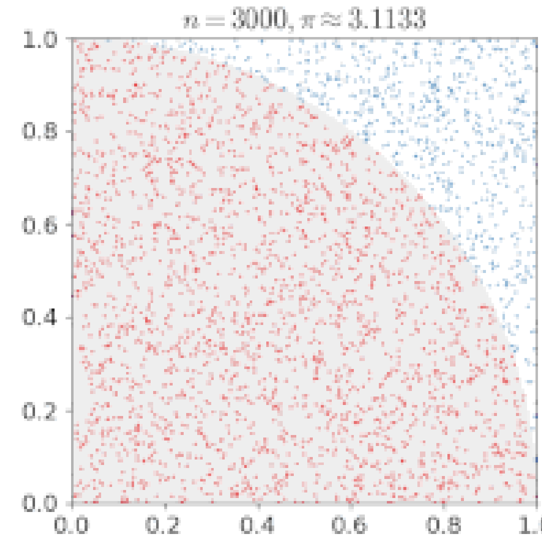
電腦模擬的實例

- 實例一、如何估計 π？(模擬積分)

$$\frac{\pi}{4} = \int_0^1 \frac{dx}{1+x^2} = \int_0^1 (1 - x^2 + x^4 - x^6 + \cdots)dx$$







https://upload.wikimedia.org/wikipedia/commons/thumb/8/84/Pi_30K.gif/330px-Pi_30K.gif

- 如果你/妳飛鏢射得很好，可以均勻地將飛鏢射在上圖的正方形中。則π 的估計值可由下式獲得：

$$\frac{\text{射在藍色區域內的飛鏢 數}}{\text{射在正方形內的飛鏢數}} = \frac{\text{藍色區域面積}}{\text{正方形面積}} = \frac{\pi}{4}$$

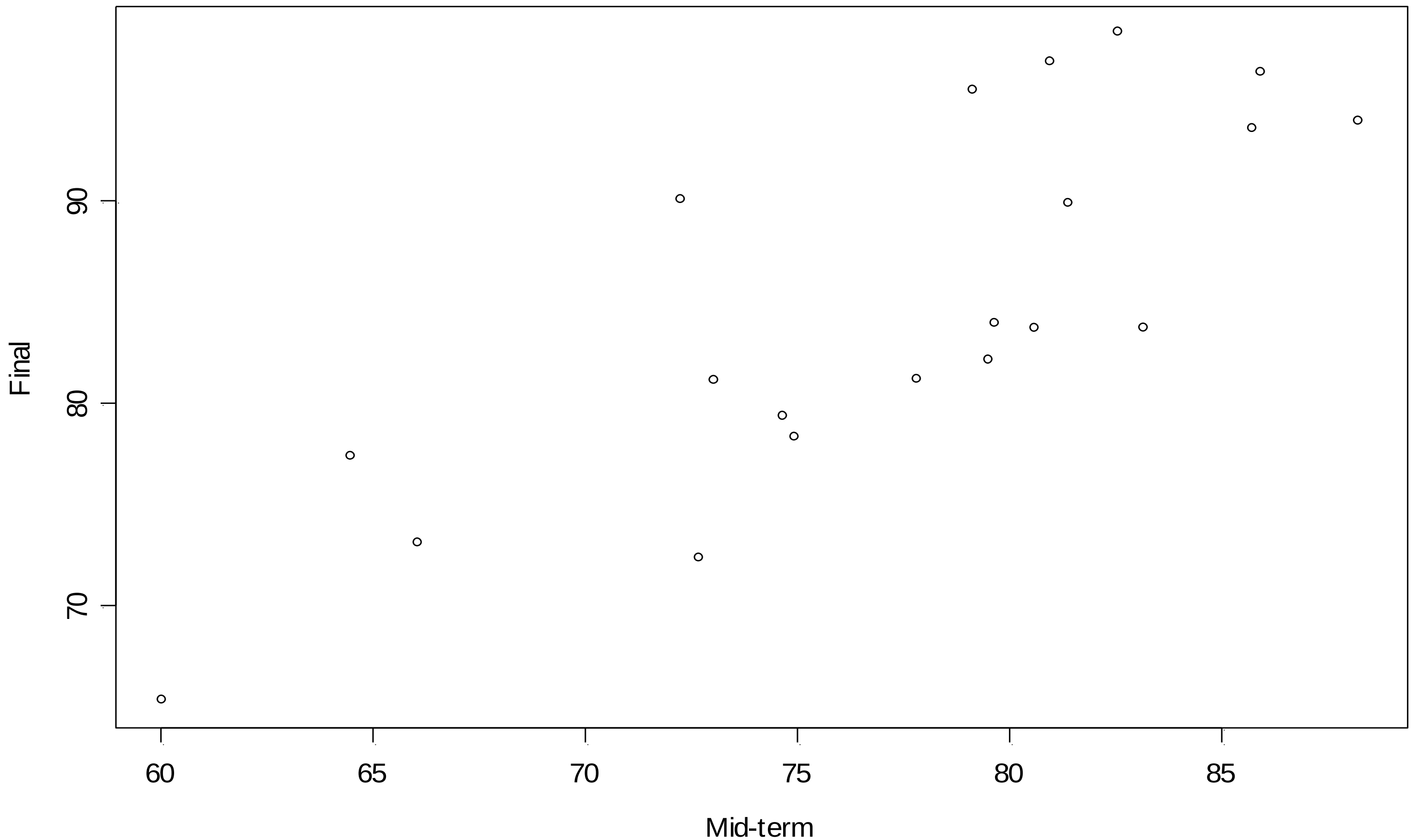
- 然而，以上述方式得出的估計值，其變異數與樣本數(模擬次數成反比)，類似問卷調查的封閉問卷，如果要得出高精確(較小的信賴區間)，需要許多次模擬？

→舉例而言，如果希望估計值的標準誤小於0.01%，則需要約1千7百萬次模擬。

- 假設有目標值的其他參考資訊，只需要約1萬次模擬即可達到上述的精確性。
  (Variance Reduction)

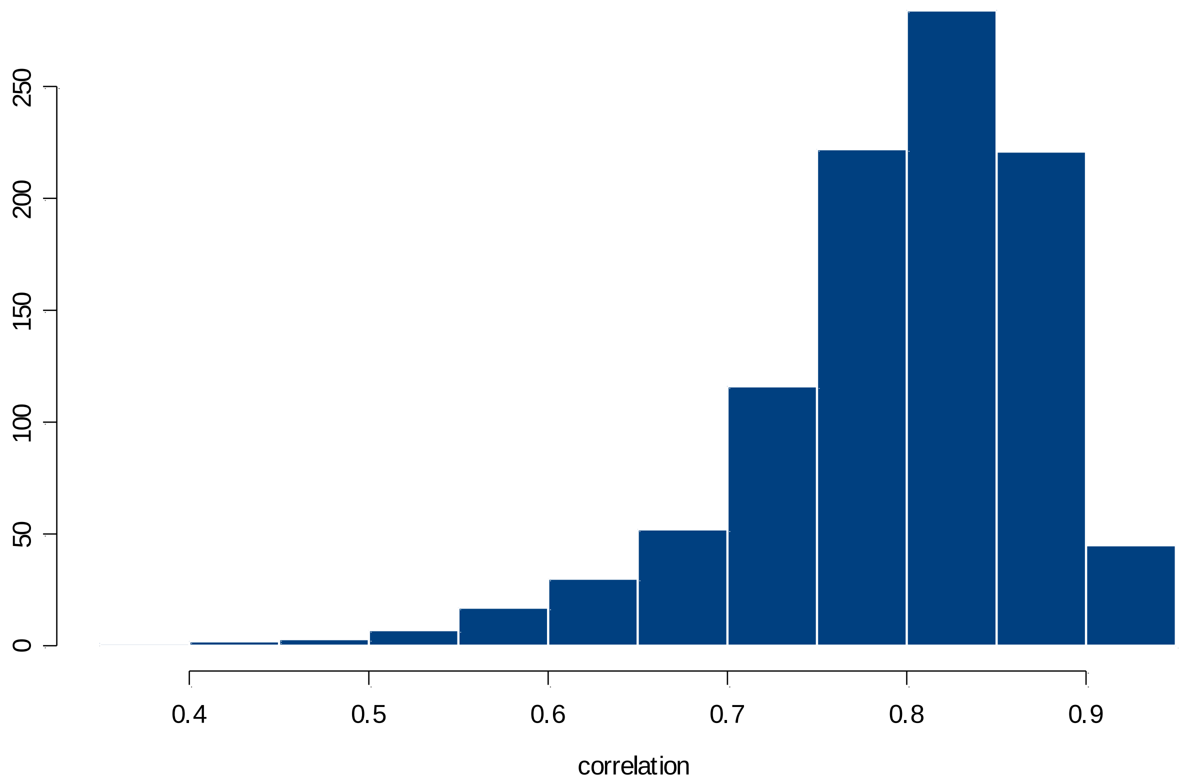
*問題：模擬積分與數值積分(Numerical Integration)有何不同？

- 實例二、我想瞭解統計學期中考與期末考兩次考試的相關性，隨機在500名學生中抽出20名學生。(Sample-resample)



- 樣本相關係數在多數軟體中都有 (約為0.8036)，但相關係數的標準差呢？

→ 拔靴法(Bootstrap)的1000次模擬計算可得標準差大約是0.0820。



- 實例三 、蒐集兩組樣本數較少的資料，如果資料不滿足常態分配，如何確定兩組期望值是否相同。(Monte Carlo p-value)

→無母數方法的檢力(power)通常比較小，且能探討的參數較有限(e.g.中位數、變異程度)，加上不見得存在適合的無母數方法(或是很難獲得精確的p-value)。

→當資料量較少時，可採用排列檢定(Permutation test)，將資料的所有可能組合列出，看看觀察值是否為極端值。

- 實例三(續)、如果兩組資料為：

  –3.58 4.93 5.27 7.24 8.33

  5.08 8.85 9.03 10.28 11.87

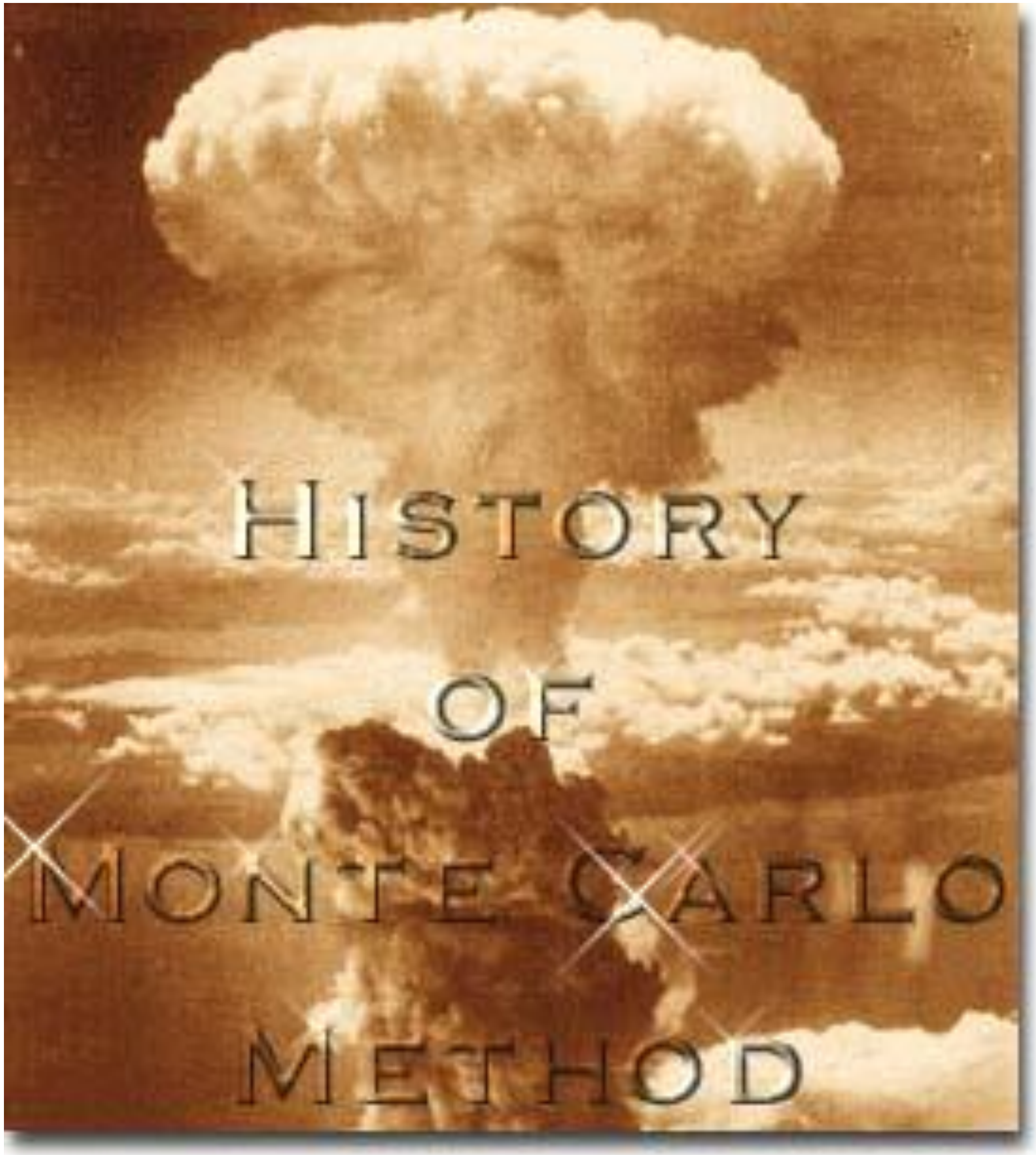
  這兩組資料是否存在差異？(註：這兩組資料由Cauchy(5)及Cauchy(10)抽出。)

→如果以 t-test 比較，結果不顯著；Kruskall Wallis 檢定p-value為0.06，但檢定的對象是中位數。

→以排列檢定而言，只有5種組合使得兩組的平均數不小於觀察值(5/252 ≈ 0.02)。

註：R的ks.test可檢查是否符合Cauchy分配，可否用於檢定、使用什麼原理？

蒙地卡羅方法的歷史



蒙地卡羅法的歷史(續)

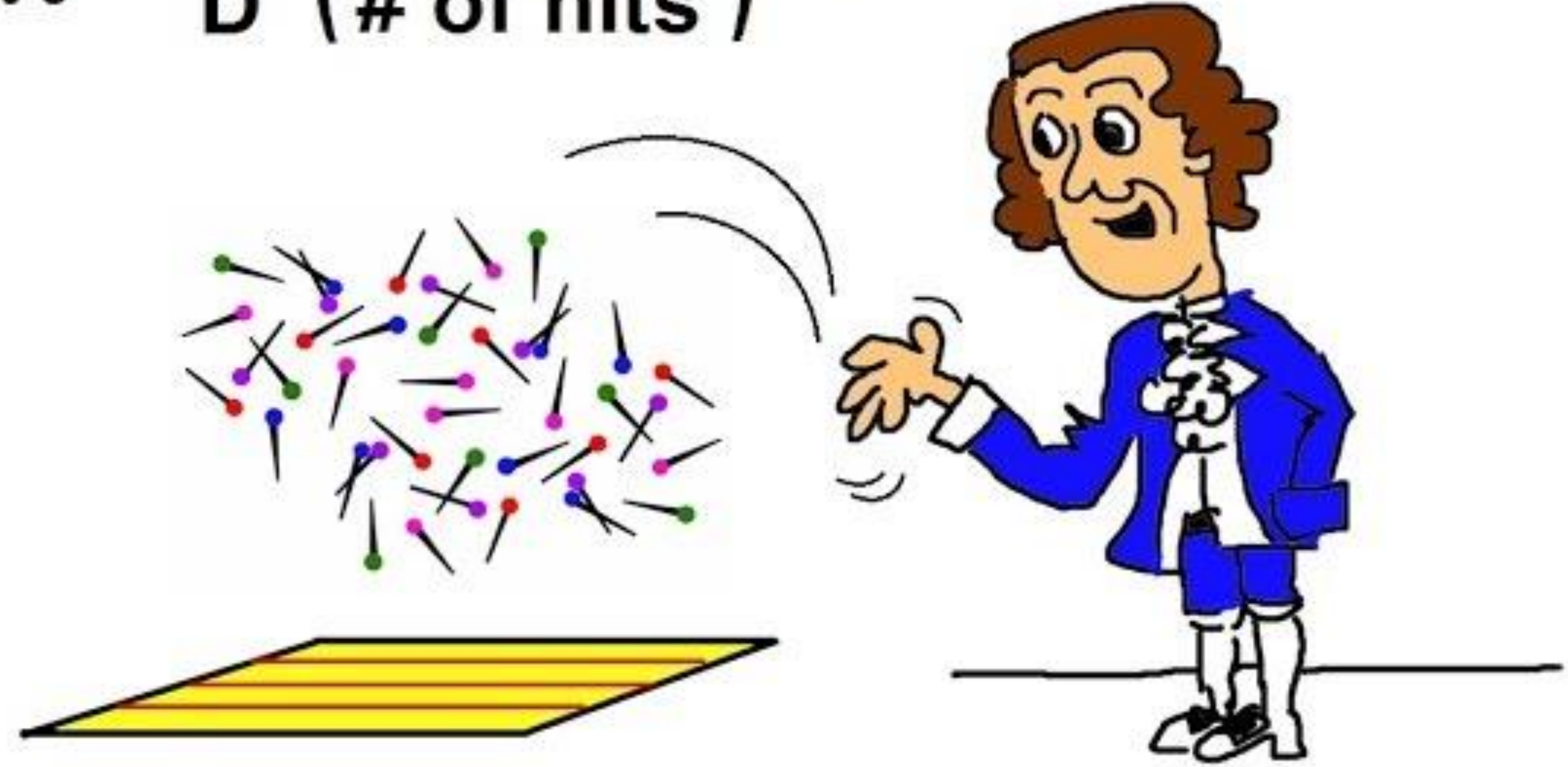
- How did Monte Carlo simulation get its name?
- The name and the systematic development of Monte Carlo methods dates from about 1940's.
- There are however a number of isolated and undeveloped instances on much earlier occasions.

蒙地卡羅法的歷史(續)

- In the second half of the nineteenth century a number of people performed experiments, in which they threw a needle in a haphazard manner onto a board ruled with parallel straight lines and inferred the value of PI =3.14… from observations of the number of intersections between needle and lines.
- In 1899 Lord Rayleigh showed that a one-dimensional random walk without absorbing barriers could provide an approximate solution to a parabolic differential equation.

Buffon's Needles and Calculating π

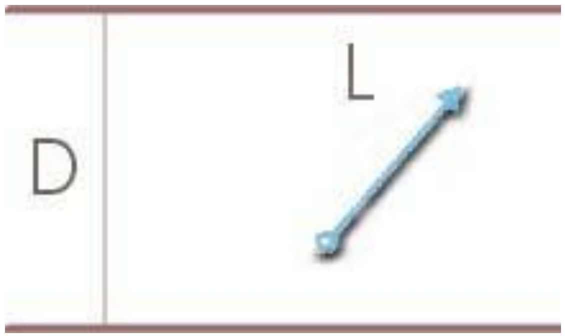
$$\pi = \frac{2L}{D}\left(\frac{\#\text{ of tries}}{\#\text{ of hits}}\right) = 3.1415...$$



https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DhHSWNsnJGeE&psig=AOvVaw32hsD5M7U9sPHCxCpCwxgz&ust=1613570738349000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCMDywczJ7u4CFQAAAAAdAAAAABAI

模擬的知名範例：Buffon Needle

Buffon's original form was to drop a needle of length L at random on grid of parallel lines of spacing D.



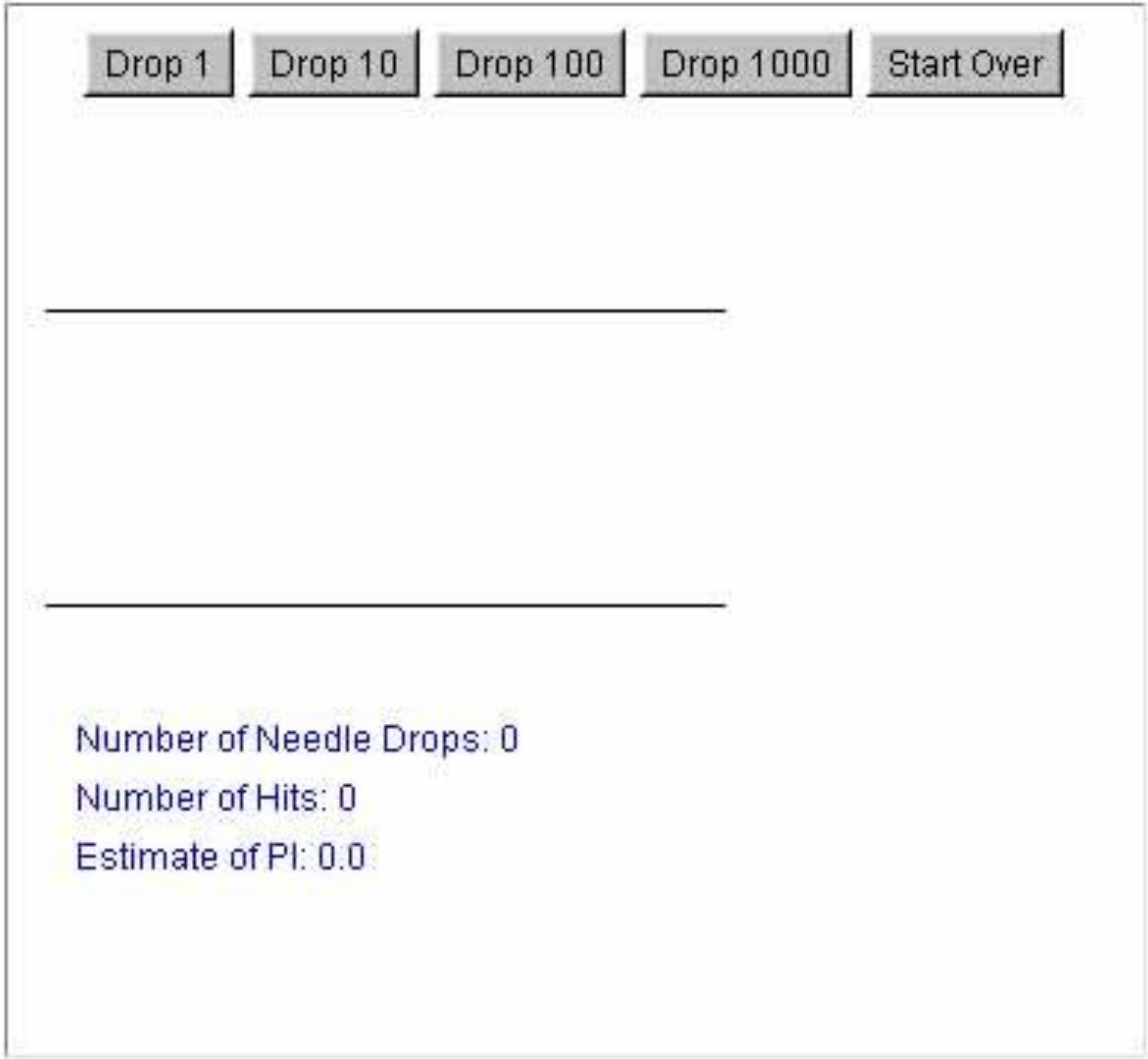
For L less than or equal D we obtain

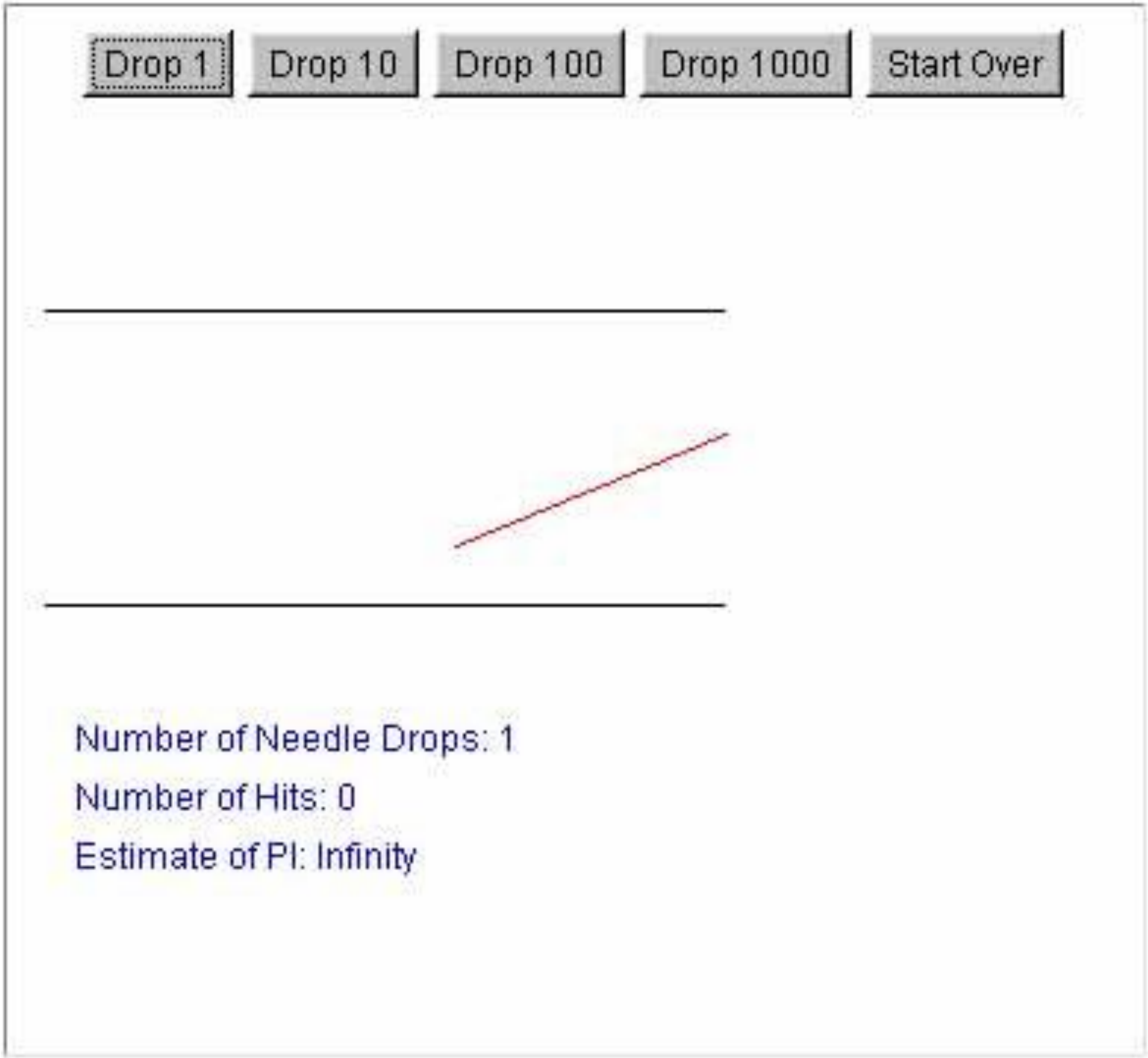
$$P(\text{needle intersects the grid}) = 2 \cdot L / PI \cdot D.$$

If we drop the needle N times and count R intersections we obtain

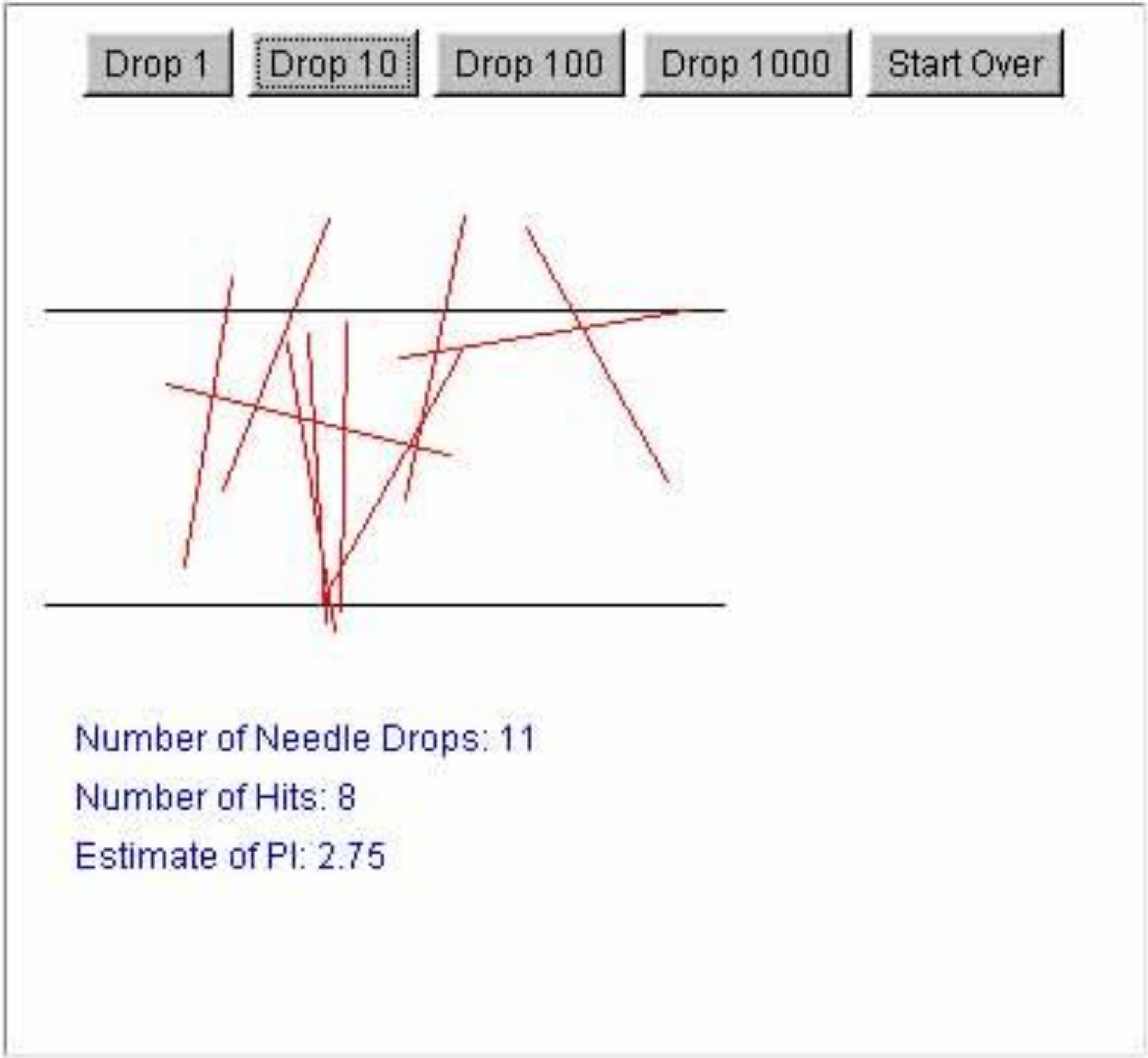
$$P = R / N,$$

$$PI = 2 \cdot L \cdot N / R \cdot D.$$

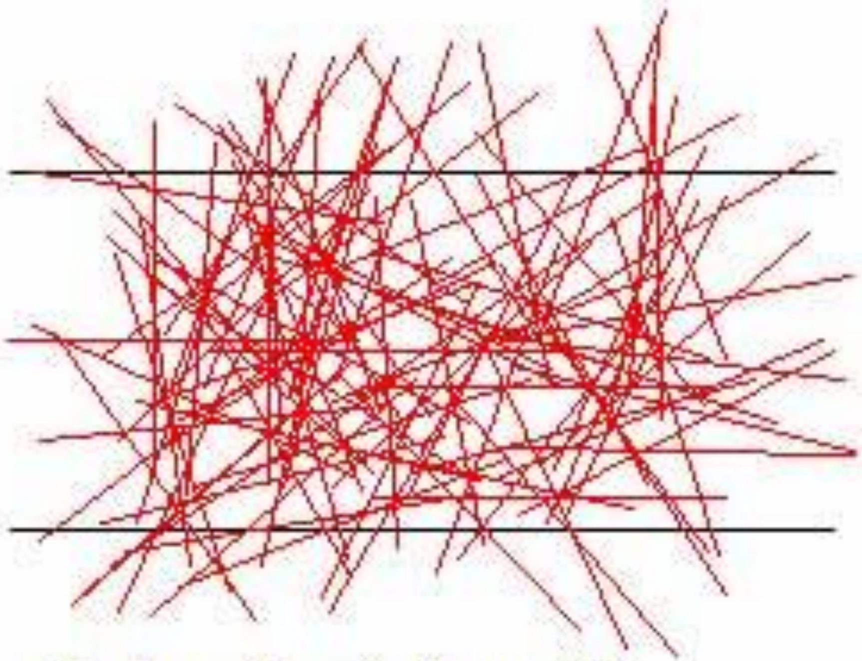

Drop 1
Drop 10
Drop 100
Drop 1000
Start Over
Number of Needle Drops: 0
Number of Hits: 0
Estimate of PI: 0.0



Drop 1
Drop 10
Drop 100
Drop 1000
Start Over
Number of Needle Drops: 1
Number of Hits: 0
Estimate of PI: Infinity


Drop 1
Drop 10
Drop 100
Drop 1000
Start Over
Number of Needle Drops: 11
Number of Hits: 8
Estimate of PI: 2.75

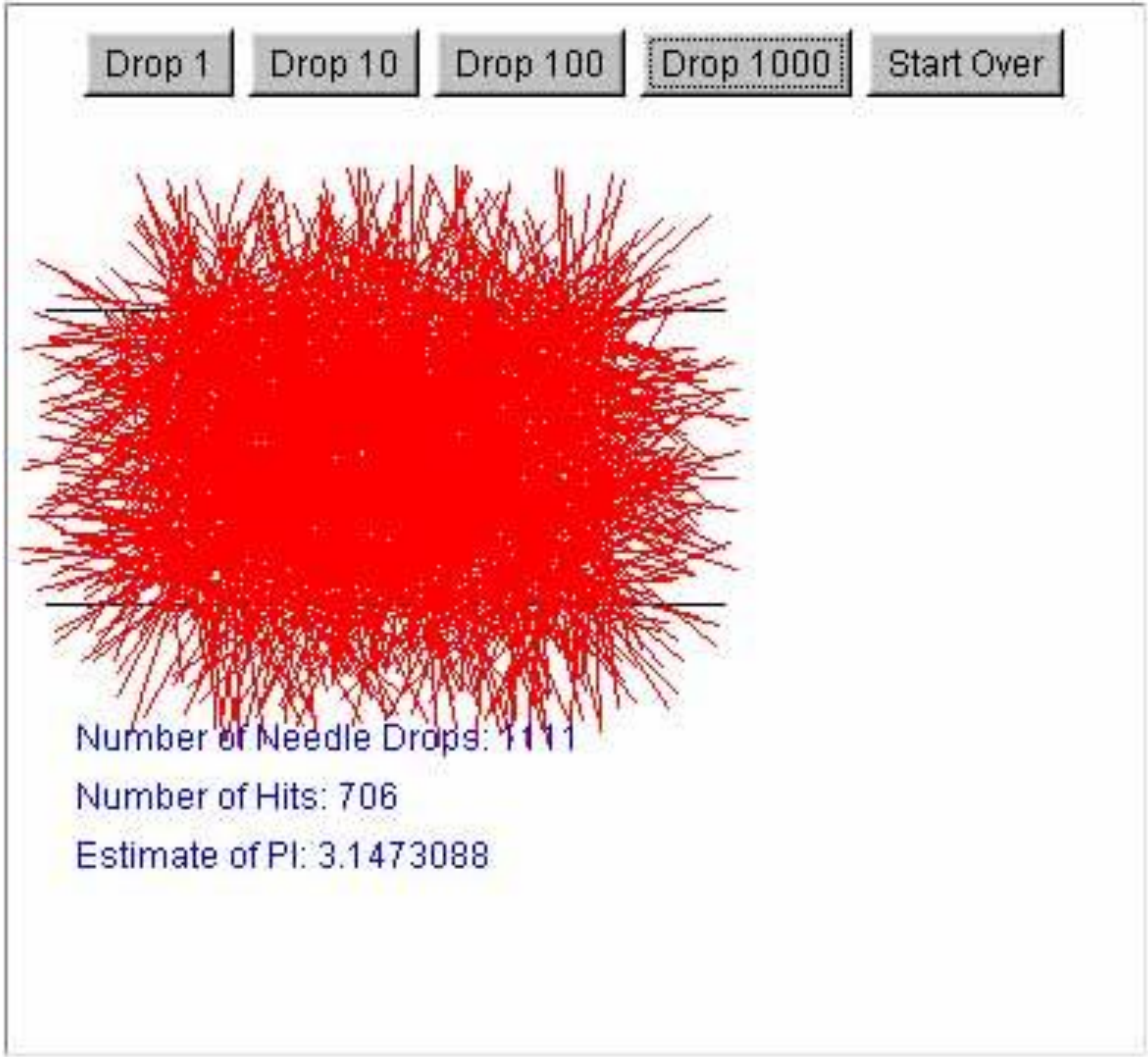
Drop 1 | Drop 10 | Drop 100 | Drop 1000 | Start Over



Number of Needle Drops: 111

Number of Hits: 69

Estimate of PI: 3.2173913


Drop 1
Drop 10
Drop 100
Drop 1000
Start Over
Number of Needle Drops: 1111
Number of Hits: 706
Estimate of PI: 3.1473088

蒙地卡羅法的歷史(續)

- In early part of the twentieth century, British statistical schools indulged in a fair amount of unsophisticated Monte Carlo work.
- In 1908 Student (W.S. Gosset) used experimental sampling to help him towards his discovery of the distribution of the correlation coefficient.
- In the same year Student also used sampling to bolster his faith in his so-called t-distribution, which he had derived by a somewhat shaky and incomplete theoretical analysis.

History of Monte Carlo Method

Student - William Sealy Gosset (1876 - 1937)

This birth-and-death process is suffering from labor pains; it will be the death of me yet. (Student Sayings)

A. N. Kolmogorov (1903-1987)

In 1931 Kolmogorov showed the relationship between Markov stochastic processes and certain integro-differential equations.

蒙地卡羅法的歷史(續)

- The real use of Monte Carlo methods as a research tool stems from work on the atomic bomb during the second world war.
- This work involved a direct simulation of the probabilistic problems concerned with random neutron diffusion in fissile material; but even at an early stage of these investigations, von Neumann and Ulam refined this particular “Russian roulette” and “splitting” methods. However, the systematic development of these ideas had to await the work of Harris and Herman Kahn in 1948.
- About 1948 Fermi, Metropolis, and Ulam obtained Monte Carlo estimates for the eigenvalues of Schrodinger equation.

John von Neumann (1903-1957)

蒙地卡羅法的歷史(續)

- In about 1970, the newly developing theory of computational complexity began to provide a more precise and persuasive rationale for employing the Mont Carlo method.
- Karp (1985) shows this property for estimating reliability in a planar multi-terminal network with randomly failing edges.
- Dyer (1989) establish it for estimating the volume of a convex body in M-dimensional Euclidean space.
- Broder (1986) and Jerrum and Sinclair (1988) establish the property for estimating the permanent of a matrix or, equivalently, the number of perfect matchings in a bipartite graph.

亂數(Random Number)的產生

- 早期的亂數多藉由實體(Physical)方法產生，例如：擲銅板、骰子、撲克牌、輪盤。

→缺點：速度慢；無法複製。

- 之後(1955)也有Rand公司的百萬個數字組成的亂數表，亂數可複製。

→缺點：速度慢；若模擬次數較多，亂數表很容易不敷使用。

— John Von Neumann (Mid-Square Method)

e.g. 任選一個四位數的數字，計算其平方數

7033 → 49**4630**89 (取中間四位數再平方)

→ 21**4369**00

→ 19**0881**61

→ 00**7761**61

註：我們也可取六位數、二位數、…

e.g. 33 → 1**08**9 → **6**4 not good！

123456 →**241**383 → **265**752

(but small numbers → small numbers)

何謂亂數？

- 上述的方法產生的數字是否為亂數？

→只能稱為假亂數(Pseudo-random or Quasi-random numbers)。

- Pseudo-random: A sequence of deterministic numbers which have the same relevant statistical properties of random number.

→這些「亂數」其實是藉由某種函數而產生的(不含隨機性)固定數。

- **對於「亂數」的要求**

  假設目標是$\{1,2,...,m\}$的亂數，則希望產生的數值能滿足以下三個要求：

  1. Uniformly distributed

  2. Statistically independent

  3. Reproducible

- 問題：如何驗證產生的數值滿足以上三個條件？

線性同餘法 (Linear Congruential Method; LCG)

- 同餘(Congruential)法是現今最廣為使用的方法，這個方法是由Lehmer(1951)，首先提出基本原理為

$$X_{i+1} = aX_i + c \pmod{m}$$

其中a, c, m為整數。

→a: multiplier (乘數)

c: increment (增量)

m: modulus (除數)

同餘法(續)

- First step: Given X_0 (seed)

  Second step: $X_1 = aX_0 + c - mK_0$

  … …

→例如：$a = c = 3, m = 5, X_0=1$

$\Rightarrow X_i = 3, 2, 4, 0, 3, \cdots$ (週期為4)

$\Rightarrow X_i = 1,1,1,1,1, \cdots$ (週期為1)

註：Seed的選擇可能影響亂數的產生。

同餘法(續)

- 基本要求：

→週期愈長愈佳、數字出現無特殊規律、Seed的影響儘量減少。

- 不錯的亂數選擇範例：

→ 十進位：$X_{i+1} = 101 X_i + 1 \pmod{1000}$

(但週期稍微短了些！)

→ 二進位：$X_{i+1} = (2^{17} + 3) X_i \pmod{2^{35}}$

註：週期短會有什麼缺點？

- 定理一、 LCG具有全週期，若且唯若：

(1) m 和 c 互質；

(2)假如質數 p 可整除 m，則 p 亦可整除 a-1；

(3)假如4可整除 m，則4亦可整除 a-1。

- 範例：$m = 16, a = 5, c = 1, X_0 = 3$

3→0→1→6→15→12→13→2→11→8→ 9→ 14→7→ 4→5→10→3→0→1→…

- 從上述的範例中，可知 X_i 與 X_{i+1} 的相關係數等於0.311，同餘法產生的亂數間似乎不見得互相獨立。

→同餘法產生的亂數 X_i 與 X_{i+1} 的相關係數間存有以下關係：

$$\rho = \frac{1}{a} - \left(\frac{6c}{am}\right)\left(1 - \frac{c}{m}\right) \pm \frac{a}{m},$$

最大值在 $a = \sqrt{m}$ 時，相關係數約為

$\frac{2}{\sqrt{m}}$，當 m 較大時很接近0。

中國剩餘定理
(Chinese Remainder Theory)

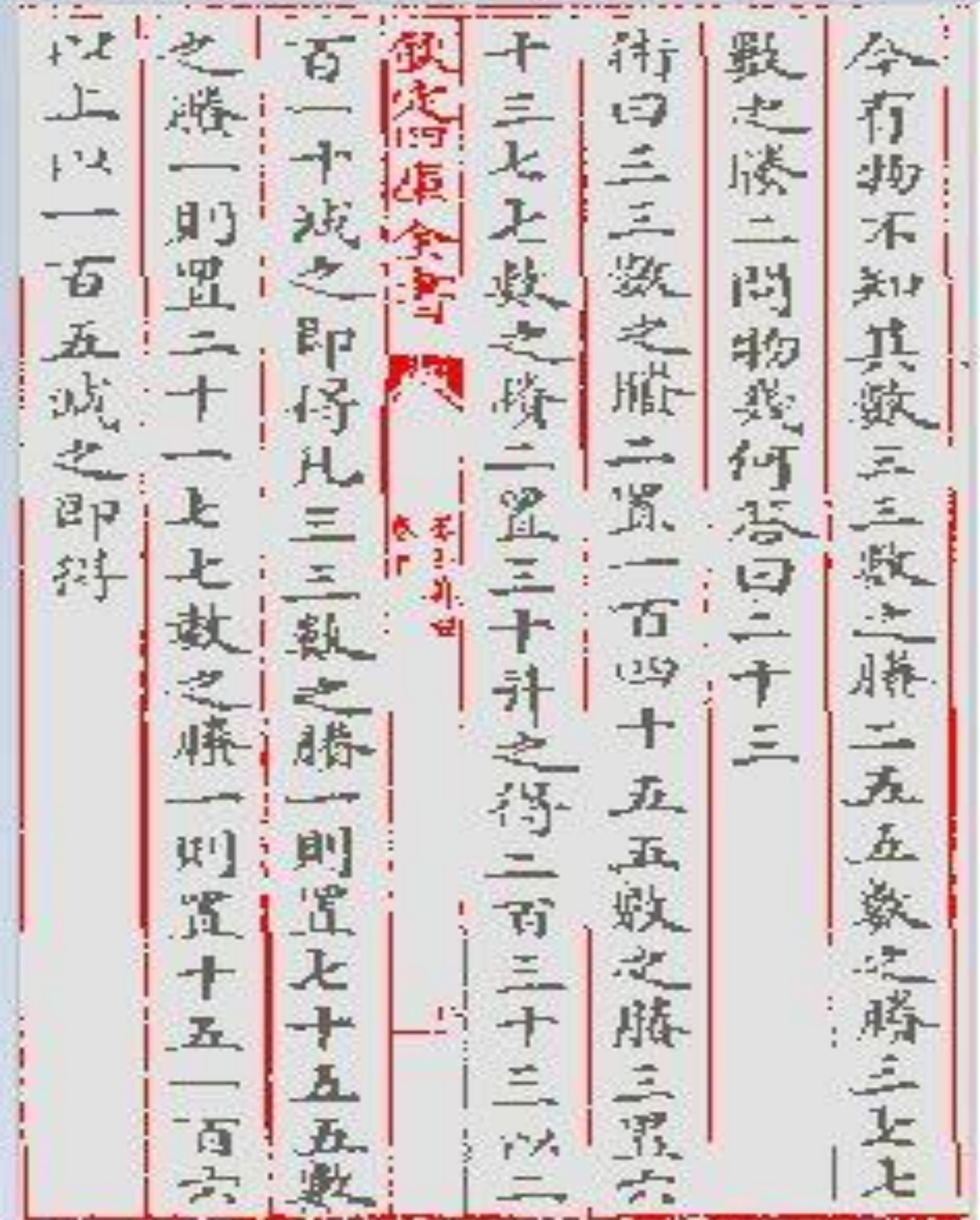
- 一元線性同餘方程組問題最早可見於中國南北朝時期（公元5世紀）的數學著作《孫子算經》卷下第二十六題，叫做「物不知數」問題，原文如下：

有物不知其數，三三數之剩二，五五數之剩三，七七數之剩二。問物幾何？

→韓信點兵？

→類似範例：

三人同行七十希，五樹梅花廿一支，七子團圓正半月，除百零五使得知。



今有物不知其數三三數之賸二五五數之賸三七七
數之賸二問物幾何答曰二十三
術曰三三數之賸二置一百四十五五數之賸三置六
十三七七數之賸二置三十并之得二百三十三以二

百一十減之即得凡三三數之賸一則置七十五五數
之賸一則置二十一七七數之賸一則置十五一百六
以上以一百五減之即得

從韓信點兵到拉格朗日插值法

韓信點兵：找個數除 3 餘 2，除 5 餘 3，除 7 餘 2。

除數	餘數	餘數	餘數	餘數	餘數	餘數	餘數
3	2	0	1	2	0	0	2
5	3	0	1	0	3	0	3
7	2	0	1	0	0	2	2
		105k	105k+1	5×7	$3\times 3\times 7$	$2\times 3\times 5$	$128+105k$
				35	63	30	$30+63+35$

拉格朗日：找多項式 $f(x)$ 滿足 $f(1)=-2$，除 $f(2)=-1$，除 $f(3)=2$。
找 $f(x)$ 除以 $x-1$ 餘 -2，除以 $x-2$ 餘 -1，除以 $x-3$ 餘 2。

除式	餘數	餘數	餘數	餘數
$x-1$	-2	-2	0	0
$x-2$	-1	0	-1	0
$x-3$	2	0	0	[illegible]
		$-2\dfrac{(x-2)(x-3)}{(1-2)(1-3)}$	$-1\dfrac{(x-1)(x-3)}{(2-1)(2-3)}$	

https://i.ytimg.com/vi/lrEOt2nCnBE/maxresdefault.jpg

- 組合產生器：(多數的軟體採用的方式)

→組合產生器是利用兩個或兩個以上的亂數產生器，來組合成一組新的亂數產生器。例如(Wichmann and Hill, 1982)：

$$x_i = 171 x_{i-1} (\text{mod} 30269)$$

$$y_i = 172 y_{i-1} (\text{mod} 30307)$$

$$z_i = 170 z_{i-1} (\text{mod} 30323)$$

$$u_i = \frac{x_i}{30269} + \frac{y_i}{30307} + \frac{z_i}{30323} \quad (\text{mod} 1)$$

同餘法(續)

- 由同餘法(或其他方法)產生介於 0 至 $m-1$ 的亂數，除以 m 後即成為0與1間的亂數，也就是類似 $U(0,1)$ 的隨機變數。

→其他分配的亂數可藉由 $U(0,1)$ 產生。

(例如：$Exp(1) = -\log(U(0,1))$。)

- 掌上型計算機(諸如Casio)使用的也是同餘法：

$$U_{n+1} = (\pi + U_n)^5 \pmod{1}$$

乘法型LCG (Multiplicative Generator)

- 為減少計算量，LCG中令 $c = 0$ 稱為乘法型LCG。

$$X_{i+1} = aX_i \pmod{m}$$

→缺點：無法達到全週期。

→若選擇的初始值 X_0 與 m 互質，再加上其他條件，可達到最大週期(半週期)。SAS使用的亂數產生器：

$$X_{i+1} = 397{,}204{,}094 X_i \pmod{2^{31} - 1}$$

幾個知名的乘法型LCG

- 除了SAS外，MATLAB在第5版前也使用乘法型LCG，週期可達 $2^{32}-1$：

$$X_{i+1} = 16{,}807 X_i \,(\bmod\, 2^{31}-1)$$

- Vax在1993年前使用下列產生器，雖然週期可達2^{32}，但獨立性檢定的效果不佳：

$$X_{i+1} = 69{,}069 X_i \,(\bmod\, 2^{32})$$

註：獨立性檢定可參考Marsaglia (1993), *"Monkey tests for random number generator."*

移動暫存產生器 (shift-register generator)

- Tausworthe在1965年所發展而成的，因此也稱為Tausworthe產生器，移動暫存產生器的設計類似密碼傳遞方法，都是直接用位元(bit)值形成亂數。例如：

$$X_{i+1} = X_i + X_{i-k} \pmod 2$$

- 著名的範例→Fibonacci Sequence

$$X_{i+1} = X_i + X_{i-k} \pmod 1$$

但效果不佳。(效果隨 k 增大而改進！)

好的亂數產生器的要件

- **長週期**：週期長則不易發生重覆使用相同亂數數列的情形。
- **良好的統計性質**：希望亂數產生器所產生的亂數數列能夠滿足統計上獨立且均勻分配的性質。
- **有效率的電腦執行**：為了便於模擬，相同的亂數數列必須具有重覆執行（reproduce）的能力。演算的速度快，亂數的產生過程中佔用較少的時間和記憶體。

- 如何確定亂數具有*U(0,1)*的特性？
  (亂數需具有哪些特質？)

→ 均勻分配(Uniformity)

多半都使用適合度(Goodness-of-fit)的檢定，確定亂數具有要求的特質，常用的方法有卡方適合度檢定(Chi-square goodness-of-fit)及Kolmogorov-Smirnov 檢定。

→ 互相獨立(Independence)

使用的多為無母數方法，例如：Gap test, Up-and-down test, run test.

常見統計軟體的亂數產生器

- 統計系常見的統計軟體：

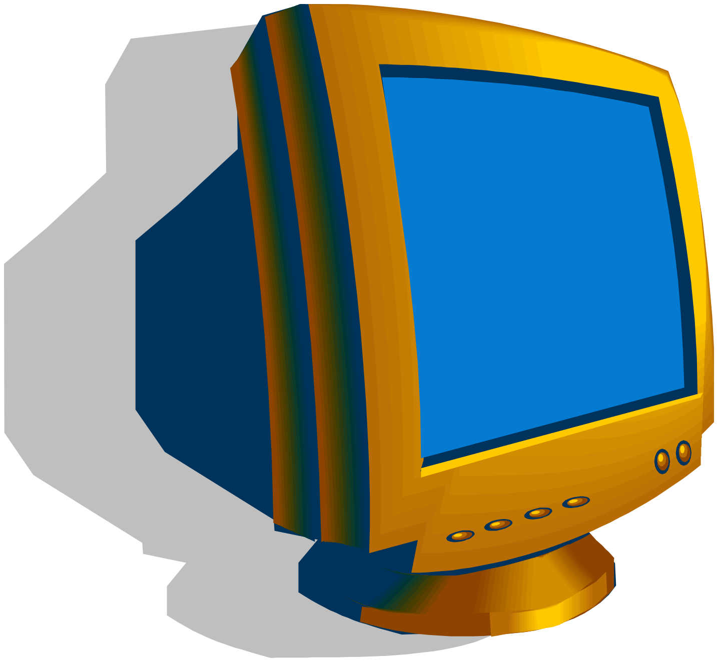
→ SAS

→ SPSS

→ S-Plus

→ Minitab

→ EXCEL



SAS的亂數產生

- SAS 6.12 內設U(0,1)亂數產生器：

→其中乘數設為397,204,094，除數設為 $2^{31}-1$，先藉由下式產生 x_i：

$$x_i = 397{,}204{,}094 x_{i-1} (\bmod 2^{31}-1) \quad i = 1,2,3,\ldots$$

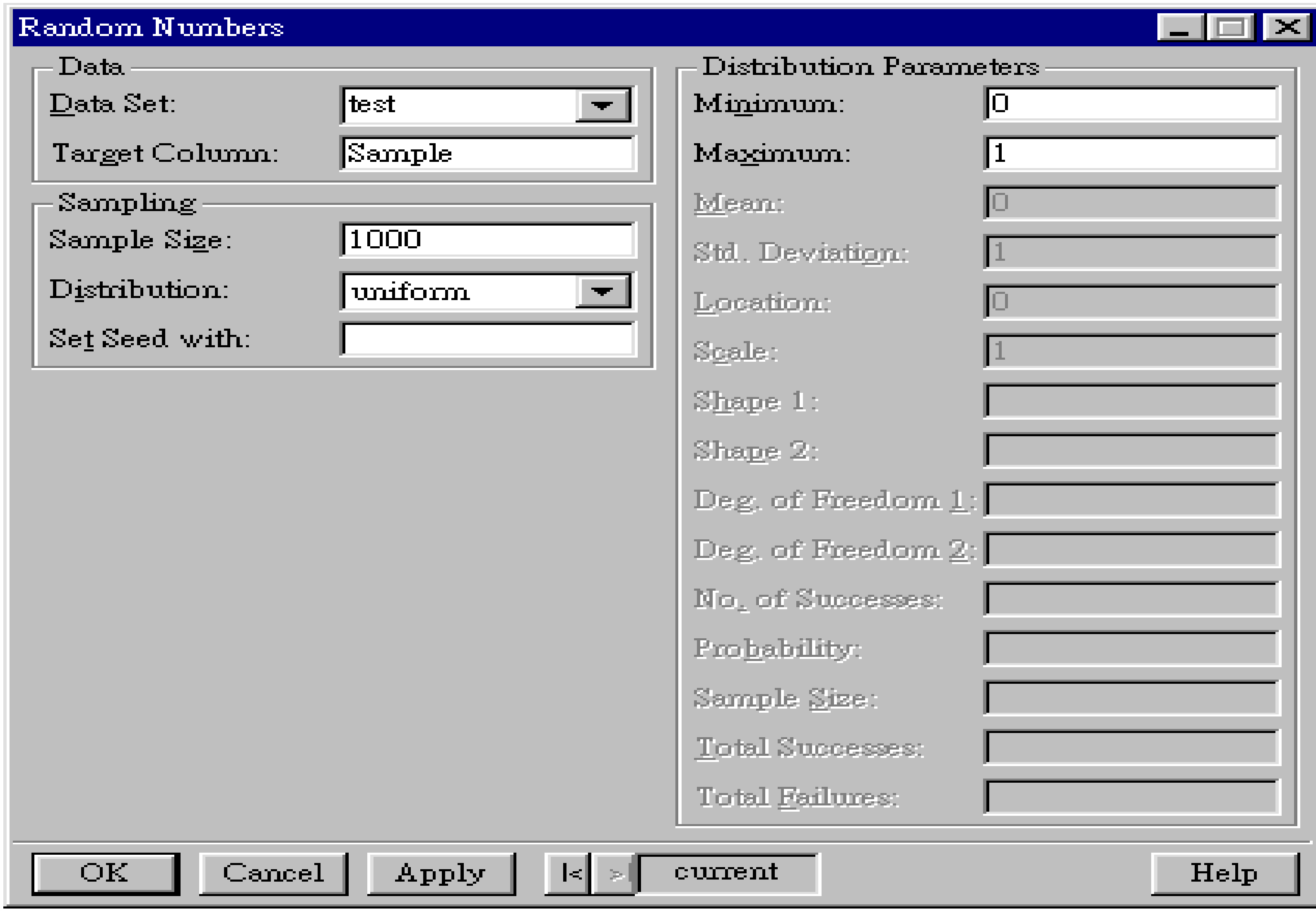
再除以 $2^{31}-1$ 而產生U(0,1)的亂數，即

$$u_i = \frac{x_i}{2^{31}-1}$$

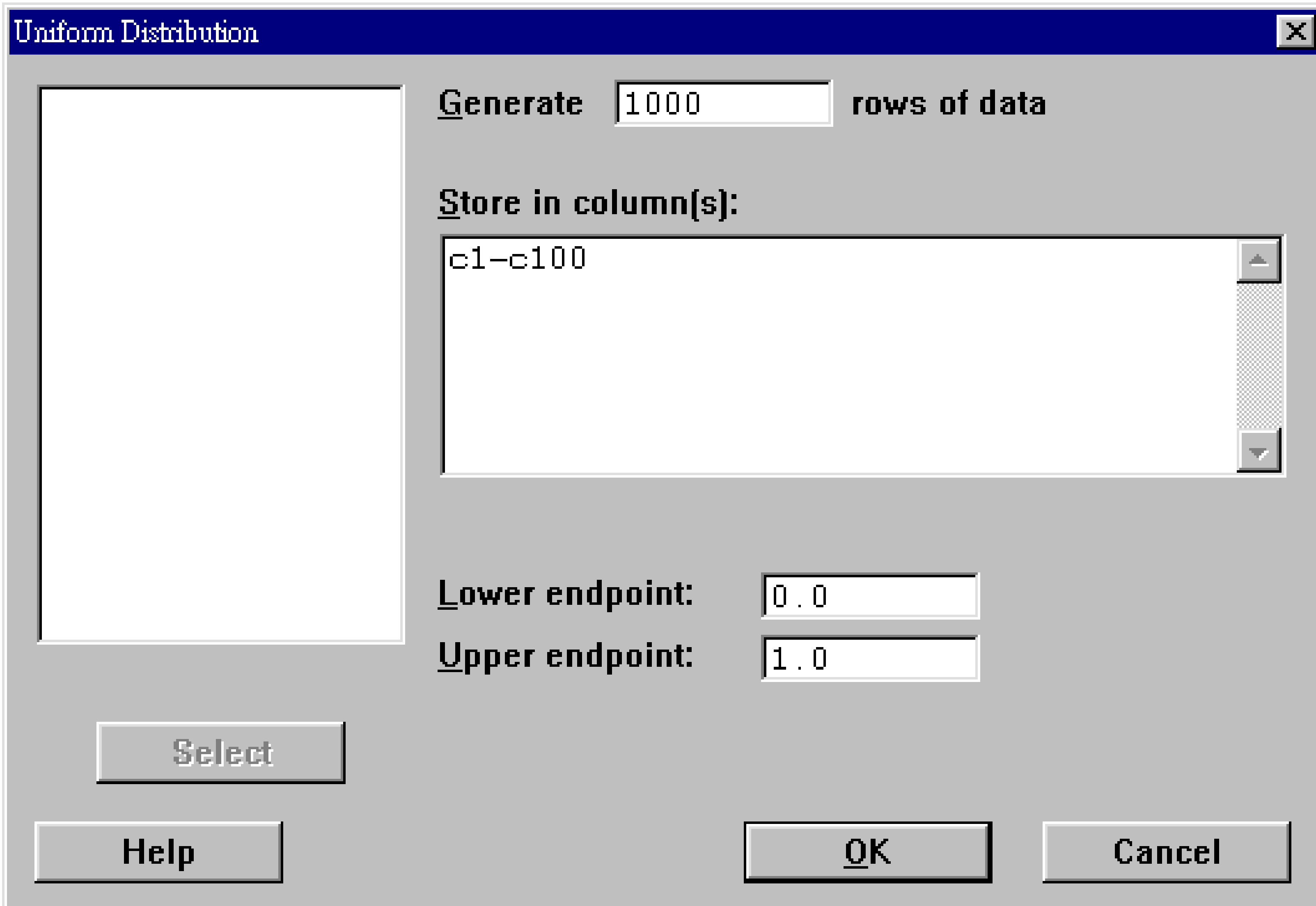
SPSS的亂數產生

計算變數
目標變數(T):
v1
=
類型註解(L)...
sn
數值運算式(E):
UNIFORM(1)
+
<
>
7
8
9
-
<=
>=
4
5
6
*
=
~=
1
2
3
/
&
|
0
.
**
~
()
刪除
函數(F):
TRUNC(數值表示式)
UNIFORM(最大值)
UPCASE(字串表示式)
VALUE(變數)
VARIANCE(數值表示式,數值
XDATE.DATE(日期值)
若(I)...
確定
貼上語法(P)
重設(R)
取消
輔助說明

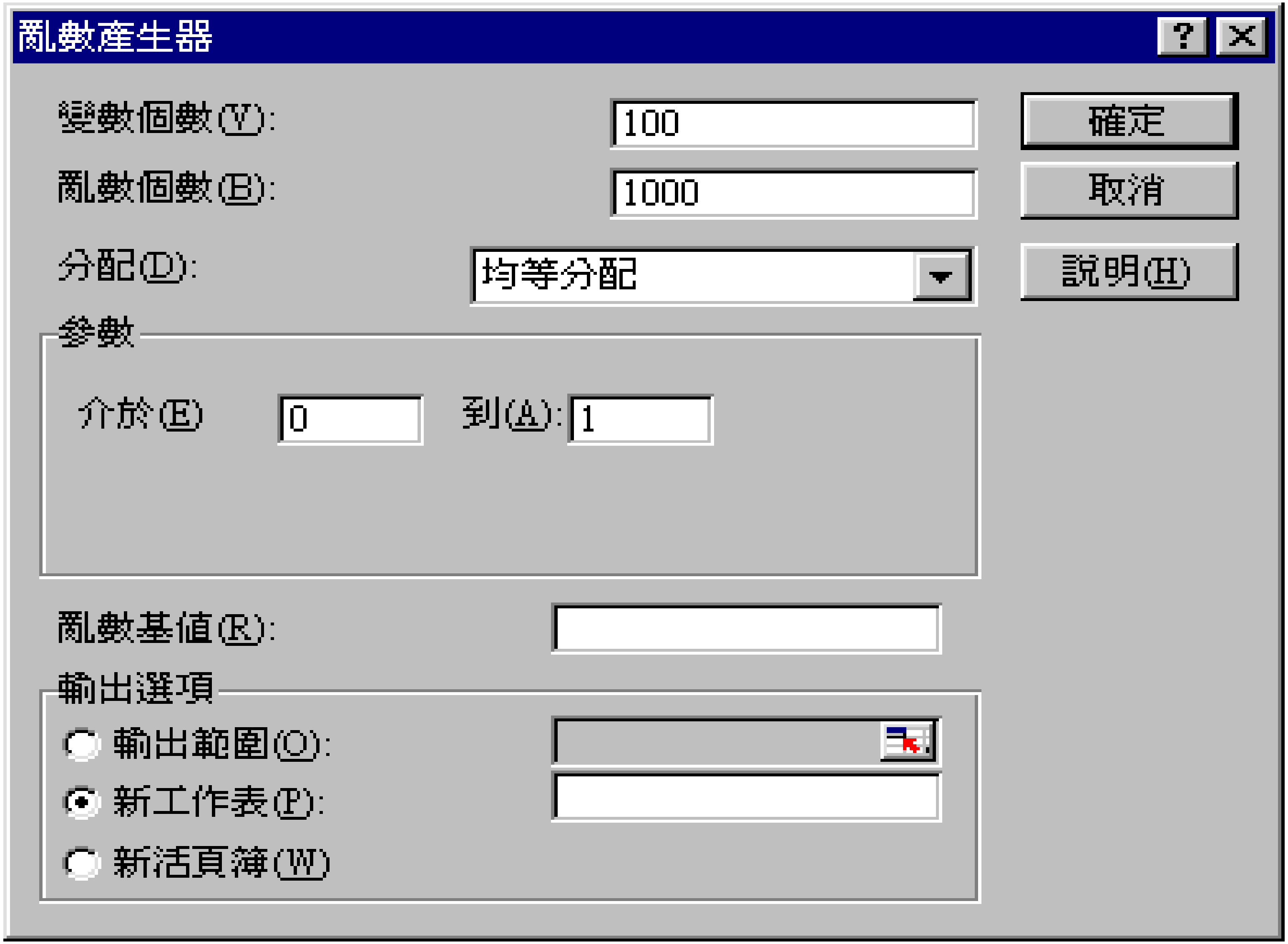
S-Plus/R的亂數產生



Minitab的亂數產生



EXCEL的亂數產生



亂數產生器
變數個數(V):
100
亂數個數(B):
1000
分配(D):
均等分配
參數
介於(E)
0
到(A):
1
亂數基值(R):
輸出選項
輸出範圍(O):
新工作表(P):
新活頁簿(W)
確定
取消
說明(H)

隨機變數的分配函數	SAS 6.12	S-PLUS 2000	MINITAB 12	EXCEL 97	SPSS 8.0
均勻(uniform)分配	*	*	*	*	*
常態(normal)分配	*	*	*	*	*
指數常態(lognormal)分配		*	*		*
指數(exponential)分配	*	*	*		*
貝他(beta)分配		*	*		*
咖嗎(gamma)分配	*	*	*		*
科西(cauchy)分配	*	*	*		*
羅吉斯(logistic)分配		*	*		*
韋伯(weibull)分配		*	*		*
F 分配		*	*		*
T 分配		*	*		*
卡方(chi-square)分配		*	*		*
Laplace 分配			*		*
Pareto 分配					*
伯努利(bernoulli)分配			*	*	*
二項(binomial)分配	*	*	*	*	*
幾何(geometric)分配		*	*		*
負二項(negbinomial)分配		*			*
卜瓦松(poisson)分配	*	*	*	*	*
超幾何(hypergeometric)分配		*	*		*

均勻檢定(Testing Uniformity)

- 卡方檢定(Chi-Square Test)：

→卡方檢定主要在處理類別資料的檢定，

$$卡方檢定: \chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i},$$

其中 o_i :觀察個數；e_i :理論個數

註：卡方檢定通常要求 $e_i \geq 5$；另外，

分組數 k 一般的選擇也不應太大。

- Kolmogorov-Smirnov Goodness-of-fit Test (適用於連續型資料)

→ $$F_N(x)=\frac{1}{N}(\#\text{ of } x_i \le x)=\frac{1}{N}\sum_{i=1}^{N} I_{(-\infty,x)}(x_i)$$

$$\therefore P(F_N(x)=\frac{k}{N})=\binom{N}{k}[F_X(x)]^k[1-F_X(x)]^{N-k}\rightarrow \text{二項分配}$$

→ 定理：

$$\lim_{N\to\infty} P[\sup_{-\infty<x<\infty} |F_N(x)-F_X(x)|>\varepsilon]=0$$

$$i.e.,\ \forall\,\varepsilon>0,\ F_N(x)\rightarrow F_X(x)\ \text{Uniforml y}$$

K-S Test (續)

→定義 $D_N \equiv \sup_{-\infty<x<\infty} |F_N(x) - F_X(x)|$

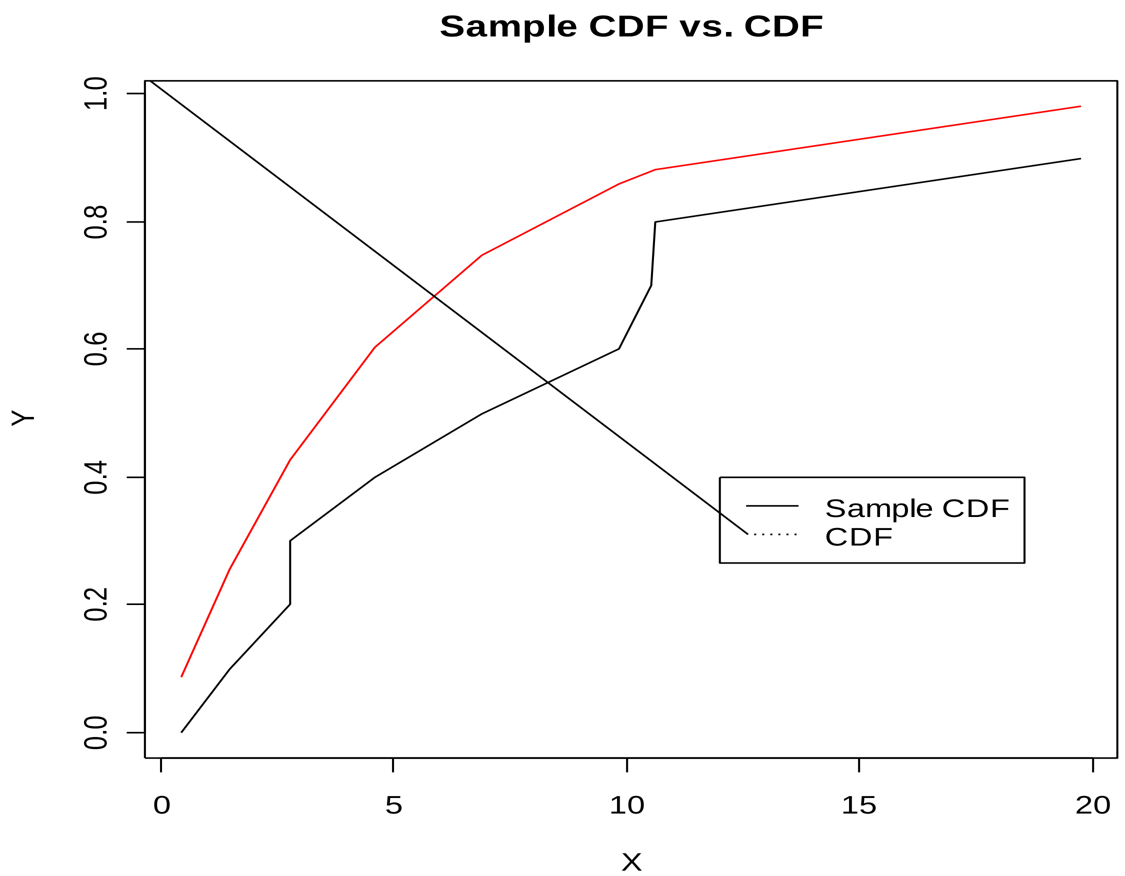
$$\Rightarrow \lim_{N\to\infty} P(\sqrt{N}D_N \le x) = [1-2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-2j^2x^2)] = H(x)$$

註：D_N 及 $H(x)$ 皆為無母數統計量。

→範例(p.156, Ross)：檢查以下十筆資料是否由*Expo(100)*抽出的亂數，66, 72, 81, 94, 112, 116, 124, 140, 145, 155。

$$D_N = 0.4831487 \quad , p\text{-value} \cong 0.012.$$

- 另一個K-S檢定的例子。隨機由Exp(λ)產生十筆觀察值，下圖是理論與樣本CDF的比較，由此可看出K-S檢定的結果。



- Chi-square 與 K-S tests 的比較：

→K-S test 只能用於連續型資料；卡方檢定則無此限制，但因卡方檢定藉由分類計算統計量，檢定結果可能因分類的個數、每個類別的定義範圍而異。

→K-S test 考慮每一觀察值，卡方檢定則分類組別計算差異。

- 對連續資料而言，兩個檢定都適用，但現今仍缺乏兩者在效率上的詳細比較。

- Chi-square 與 K-S tests 的比較(續)：

→12個 U(0,1)亂數($\sum_{i=1}^{12} U_i - 6$) 可用於近似標準常態分配，因為期望值及變異數都與標準常態分配相同。

→但近似與實際仍然不同，適合度檢定應該偵測出其中的不同(i.e., power)。

拒絕次數/1000次模擬	α＝0.01	α＝0.05	α＝0.10
χ2-test	20	93	142
KS-test	22	116	202

Cramer-von Mises Goodness-of-fit Test

→ $H_0 : F_X(x) = F_0(x)$ vs. H_1 : 存在 x 滿足 $F_X(x) \neq F_0(x)$

令 $X_{(1)},\cdots,X_{(N)}$ 為 $X_1,\cdots X_N$ 的順序統計量

$$\text{定義：} Y = \frac{1}{12N} + \sum_{i=1}^{N}[F_0(X_{(i)}) - \frac{2i-1}{2N}]^2$$

$$\because F_N(X_{(i)}) \text{ 介於 } \frac{i-1}{N} \text{ 與 } \frac{i}{N} \text{ 間} \Rightarrow \text{平均數 } \frac{2i-1}{2N}$$

→檢定表可查閱 Anderson & Darling (1952, Annals, p.193-212)；也有人稱此檢定為 Anderson-Darling test（但兩者不同）。

R軟體的常態分配檢定

- 資料是否服從常態分配較為常見，舉凡統計學提到的「經驗法則」、迴歸的殘差檢定都是其中範例。

→R也將幾種常見的常態分配（連續型）檢定組織成「nortest」，除了上述兩種檢定外，也包括「ad.test」、「lillie.test」、「sf.test」，各種方法各有特色，使用時先留意優缺點及限制。

→Q：多維常態分配如何檢定？

獨立性檢定(Independence Test)

- Gaps test:

→選取介於 0, 1間的兩個數 $0 < \alpha < \beta < 1$，

再考慮介於 α 與 β 間的亂數。介於 α 與 β 間的亂數間隔應與幾何分配有關，間隔個數

$$P(K=k)=(\beta-\alpha)\times(1-(\beta-\alpha))^k, k=0,1,\ldots$$

→再套入卡方檢定。

- Up-and-Down test：

→比較連續兩個亂數間的大小，以正號(+或1)表示遞增、負號(–或0)表示遞減。

e.g. 觀察值 0.2 0.4 0.1 0.3 0.6 0.7 0.5

→1 0 1 1 1 0 表示連續數值間的關係。

- 連續的 0 及 1 視為一個單位，稱為連(Run)，例如以上範例共有4個連。

→長度為 *k* 的連的期望值、變異數可參考 Levene and Wolfowitz (Ann. Math stat., 1944, p.58-69)，透過 *N*!種排列數推導出。

e.g. Consider the case of N=2,3,4

(1) N=2 $12 \to 1$

$21 \to 0$ Totally, 2 runs. (ave.=1)

(2) N=3 $123 \to 11$

$132 \to 10$

$213 \to 01$

$231 \to 10$

$312 \to 01$

$321 \to 00$ Totally, 10 runs. (ave.=5/3)

(3) N=4 $1234 \to 111$

$1243 \to 110$

$1324 \to 101$

$1342 \to 110$

$1423 \to 101$

$1432 \to 100$

$\dots \to \dots$ Totally, 56 runs. (ave.=7/3)

- Up-and-Down test (續)：

→連長度為一 = $(N+1)/12$、

連長度為二 = $(11N-14)/12$、

連長度為 $N-1 = 2/N!$、

連長度為 $k = \dfrac{2[(k^2+3k+1)N-(k^3+3k^2-k-4)]}{(k+3)!}$

連個數的期望值 = $(2N-1)/3$、

連個數的變異數 = $(16N-29)/90$、

$$Z = \frac{U-(2N-1)/3}{[(16N-29)/90]^{\frac{1}{2}}} \sim N(0,1)$$

- Up-and-down test (續)：

→Another so called "Run tests" in Minitab. Use a number (usually Mean) as base number and see if the observed number is greater than this number. We would expect the number of observed greater than "mean" is about half , i.e. Binomial distribution.

e.g. In Minitab, I tried 100 random number data, the following is the output：

k = 0.5419
The observed no. of runs = 61
The expected no. of runs = 50.50
55 observations above k. 45 below
The test is significant at 0.0332.

- Permutation Tests：

→將亂數依先後順序每 k 個一組，

$$(X_1, X_2, \ldots, X_k),\ (X_{k+1}, X_{k+2}, \ldots),\ \ldots$$

每組再依大小順序寫出類似 $(1,2,\ldots,k)$ 的格式，共有 k! 種可能性。

→計算每一種可能的個數，再以卡方檢定檢查每一種可能發生率是否相同。

問題：Up-and-down test 及 Permutation test 兩者適用時機有何不同？

- Coupon Collectors’ Test：

→假設亂數可分類為 k 種可能，計算每一種可能都出現時所需的樣本數。

→需要的樣本數為$\{k, k+1, k+2, \ldots\}$。

→參考Greenwood, 1955, Math comp. p.1-15

註：Permutation 及 Coupon Collectors’ tests 的共同特色是當 k 較大時，需要較大的樣本數才能進行卡方檢定。

- Q：獨立性檢定可透過x_i和x_{i+1}間的相關係數（圖形），再輔以迴歸之類的檢定？

→以R的ARIMA亂數產生器，發現相關係數存在使用限制。

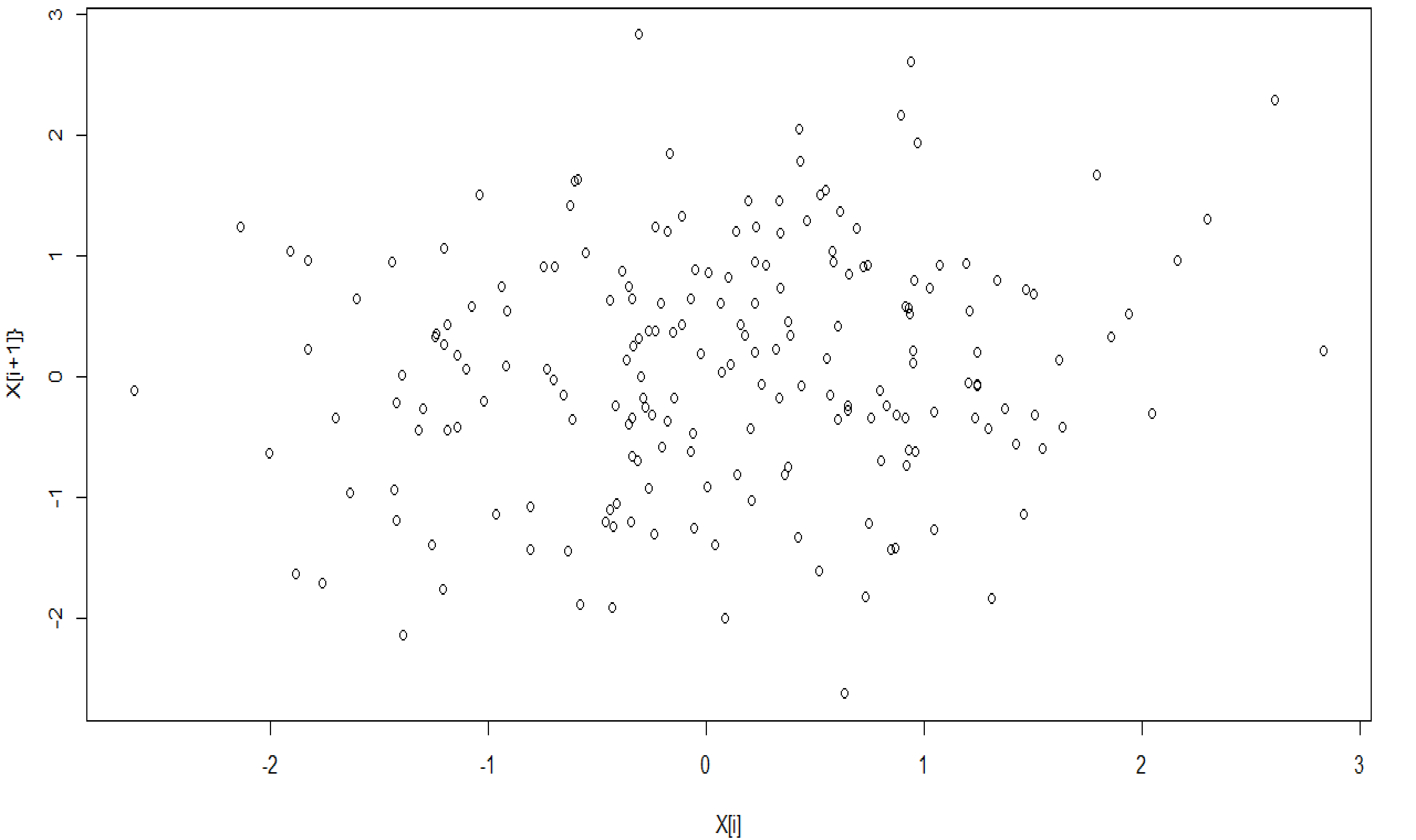
註：x_i和x_{i+k}的相關係數為k階相關。

- Q：時間數列ACF、PACF等工具檢測？

→套用這些工具都需要常態分配假設，若可透過變數轉換自然可行。

註：或可與上述獨立性檢定比較，計算不同方法的型一、型二誤差。

假設資料服從AR(2)_(0.1,0.1)



註：代入MA(2)_(0.1,0.1)亂數也有類似結果。

- Q：如何選擇獨立性檢定？

→不同類型的獨立性檢定有什麼差異，應用時如何選擇？

註：哪些與獨立有關的特性為「必要」，或是優先檢查的項目？

- 實際上並不存在「完美的亂數」，每次檢定都有型一誤差，不建議套用過多檢定，至多選擇兩、三種不同型態的獨立性檢定。

→**<u>個人觀點</u>**：根據問題需要，選擇兩種獨立性檢定。（也考慮資料多寡！）