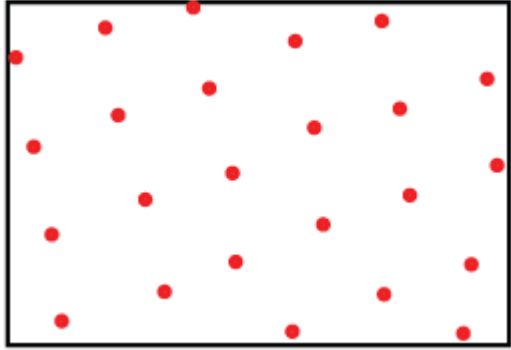


# HYPOTHESIS TESTING & POINT PATTERN ANALYSIS

計量地理學及實習 2018.11.23 杜承軒

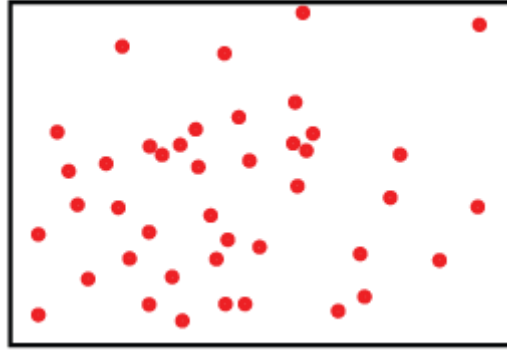
# 點型態分布

分散



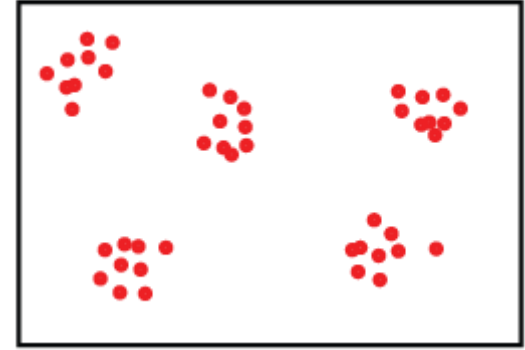
均勻分布  
uniform  
dispersion

隨機



隨機分布  
random  
independent

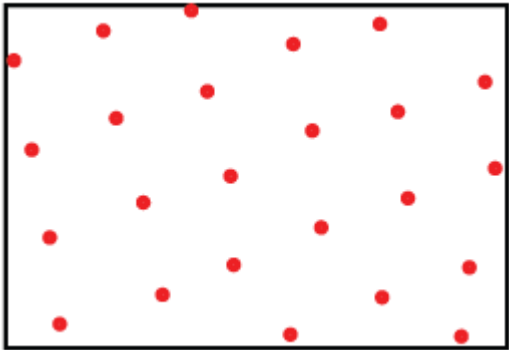
群聚



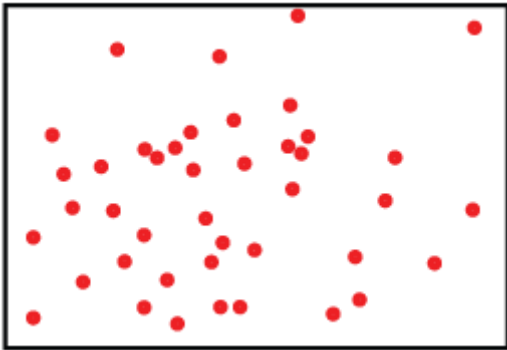
聚集分布  
cluster  
aggregated

點型態分布

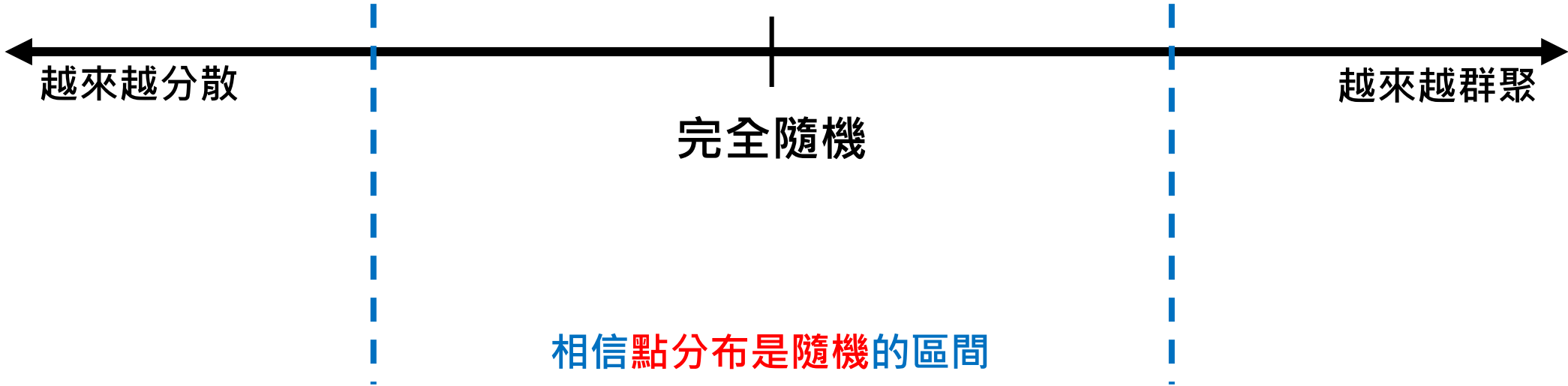
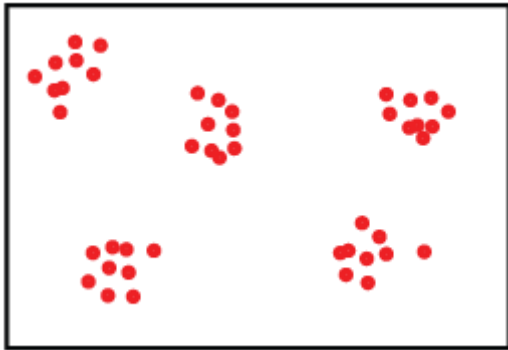
分散



隨機



群聚



## 概念複習：假設檢定

- ① 設定虛無假設與對立假設
- ② 計算統計量
- ③ 比較p值與顯著水準 $\alpha$
- ④ 決定是否拒絕虛無假設
- ⑤ 得出結論

## 概念複習：假設檢定

① 虛無假設與對立假設

虛無假設： $H_0$

現狀、沒有關係、沒有區別

② 計算統計量

③ 比較p值與 $\alpha$

對立假設： $H_a$

現狀為非、有關係、有差別

④ 拒絕虛無假設？

※通常是研究者希望證明的

⑤ 得出結論

概念複習：假設檢定

- ① 虛無假設與對立假設
- ② 計算統計量
- ③ 比較p值與 $\alpha$
- ④ 拒絕虛無假設？
- ⑤ 得出結論

雙尾檢定  
A和B有沒有區別？

$H_0$

$$\begin{aligned} \mu &= 170 \\ \mu_1 - \mu_2 &= 0 \\ s^2 &= \lambda \end{aligned}$$

$H_a$

$$\begin{aligned} \mu &\neq 170 \\ \mu_1 - \mu_2 &\neq 0 \\ s^2 &\neq \lambda \end{aligned}$$

不等於

單尾檢定  
A有沒有比B大/小？

$$\begin{aligned} \mu &= 170 & \mu &\leq 170 \\ \mu_1 - \mu_2 &= 0 & \text{或} & \mu_1 - \mu_2 \geq 0 \\ s^2 &= \lambda & & s^2 \leq \lambda \end{aligned}$$

(兩者的概念是一樣的)

$$\begin{aligned} \mu &> 170 \\ \mu_1 - \mu_2 &< 0 \\ s^2 &> \lambda \end{aligned}$$

大於、小於

## 概念複習：假設檢定

依照你的需求進行檢定，計算統計量。  
→例如：t檢定、卡方檢定.....

- ① 虛無假設與對立假設
- ② 計算統計量
- ③ 比較p值與 $\alpha$
- ④ 拒絕虛無假設？
- ⑤ 得出結論

- t檢定
  - 單一母體平均值
  - 兩母體平均差
- ANOVA 變異數分析 (F檢定)
  - 多母體平均是否一致
- 卡方檢定 ( $\chi^2$ 檢定)
  - 獨立性檢定：兩個變數是否獨立
  - 齊一性檢定：母體分配是否相同
  - 適合度檢定：樣本是否服從某機率分配  
/ 某已知關係

概念複習：假設檢定

- ① 虛無假設與對立假設
- ② 計算統計量
- ③ 比較p值與 $\alpha$
- ④ 拒絕虛無假設？
- ⑤ 得出結論

真實狀況

$H_0$  為真  
(不應拒絕 $H_0$ )

$H_0$  為假  
(應拒絕 $H_0$ )

判斷結果

拒絕 $H_0$

接受 $H_0$

型一錯誤

正確

$\alpha$ ：顯著水準，容許型一錯誤發生的機率上限

p：犯型一錯誤的機率

正確

型二錯誤

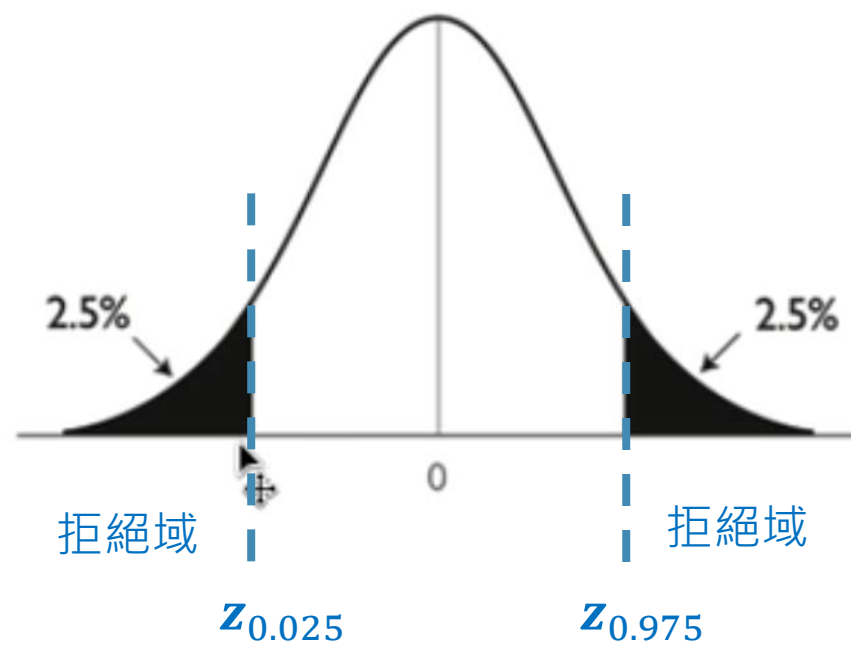


# 概念複習：假設檢定

$\alpha=0.05$

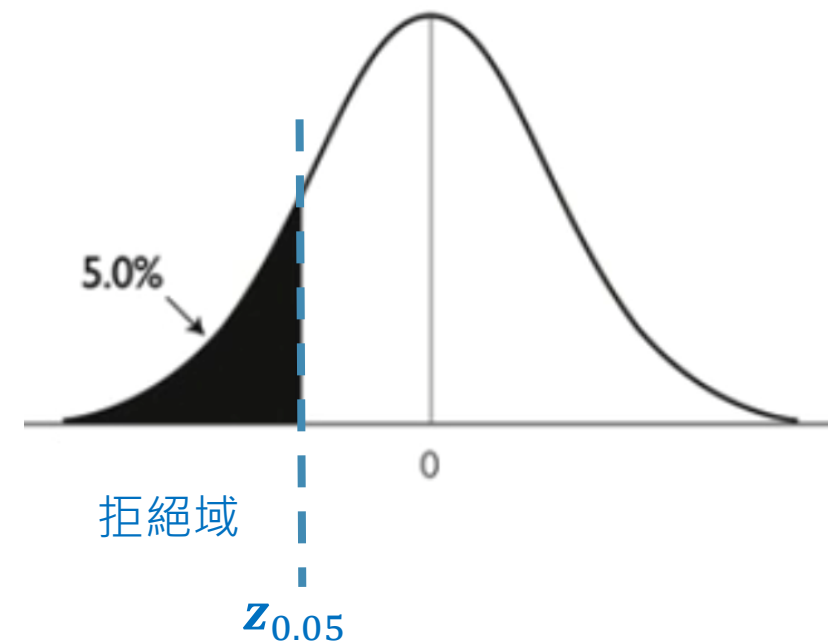
## 雙尾檢定

Two-tailed test



## 單尾檢定

One-tailed test



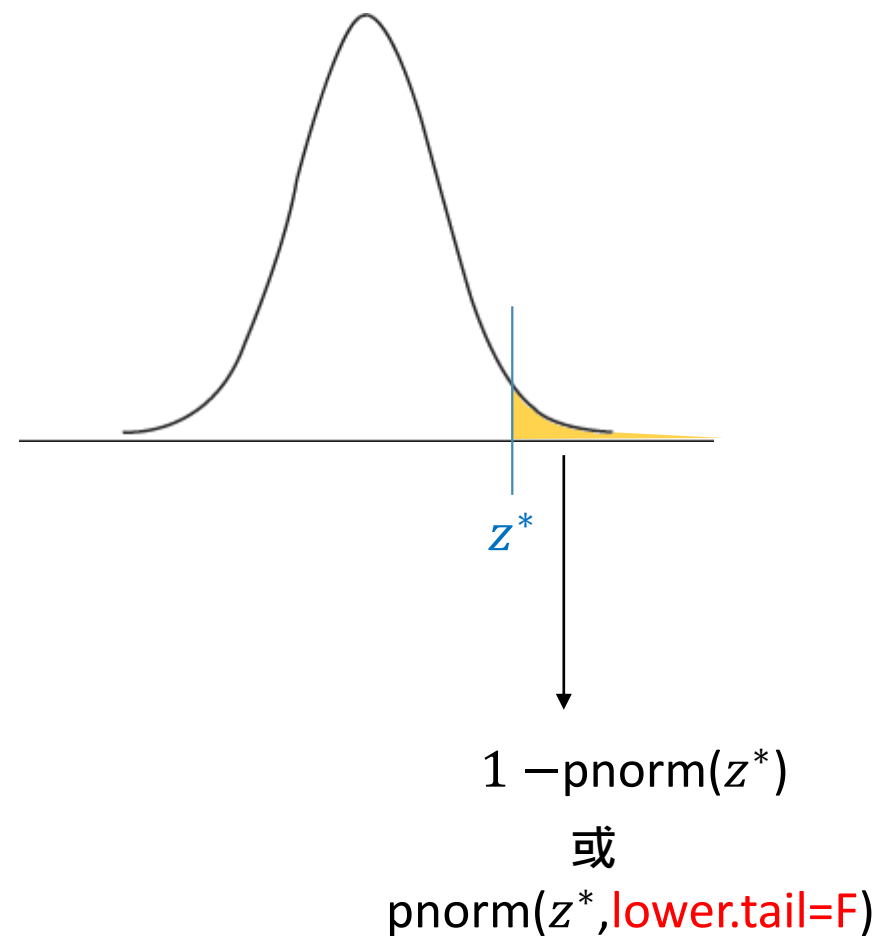
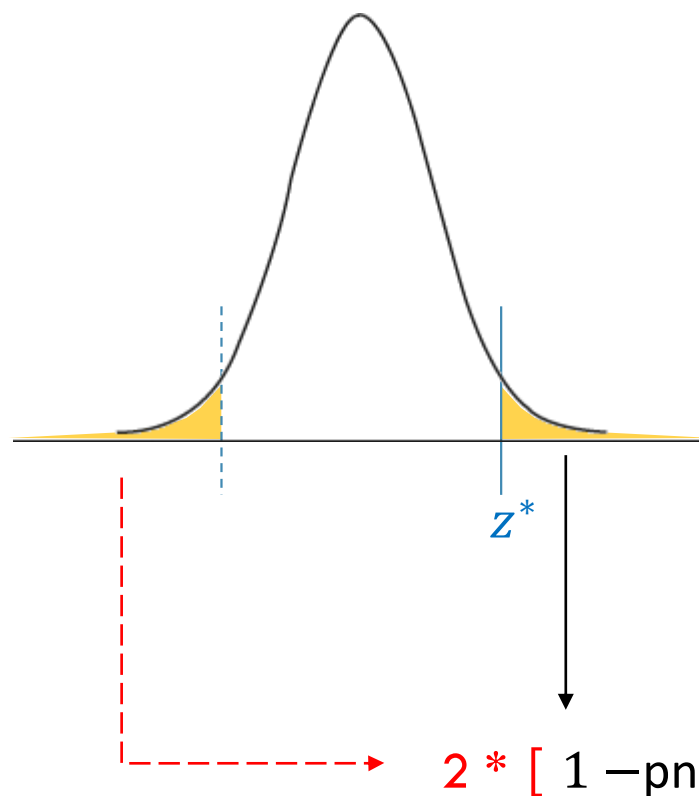
※統計量落在拒絕域→p值小於 $\alpha$ （→拒絕虛無假設）

- ① 虛無假設與對立假設
- ② 計算統計量
- ③ 比較p值與 $\alpha$
- ④ 拒絕虛無假設？
- ⑤ 得出結論

## 雙尾檢定

## 單尾檢定

- ① 虛無假設與對立假設
- ② 計算統計量
- ③ 比較p值與 $\alpha$
- ④ 拒絕虛無假設？
- ⑤ 得出結論



## 概念複習：假設檢定

① 虛無假設與對立假設

② 計算統計量

③ 比較p值與 $\alpha$

④ 拒絕虛無假設？

⑤ 得出結論

•  $p\text{值} < \alpha \rightarrow \text{拒絕} H_0$

•  $p\text{值} \geq \alpha \rightarrow \text{接受} H_0$

## 概念複習：假設檢定

① 虛無假設與對立假設

② 計算統計量

③ 比較p值與 $\alpha$

④ 拒絕虛無假設？

⑤ 得出結論

•  $p\text{值} < \alpha \rightarrow \text{拒絕} H_0$

結論：雙尾－顯著有差、有關係  
單尾－顯著較大 / 小

•  $p\text{值} \geq \alpha \rightarrow \text{接受} H_0$

結論：沒有關係、沒有區別

## 概念複習：假設檢定

$$H_0 : s^2 = \lambda \text{ (隨機分布)}$$

① 虛無假設與對立假設

② 計算統計量

③ 比較p值與 $\alpha$

④ 拒絕虛無假設？

⑤ 得出結論

•  $p\text{值} < \alpha \rightarrow \text{拒絕} H_0$

結論：雙尾－顯著有差、有關係  
單尾－顯著較大 / 小

非隨機分布

群聚現象

•  $p\text{值} \geq \alpha \rightarrow \text{接受} H_0$

結論：沒有關係、沒有區別

隨機分布

# Quadrat Analysis

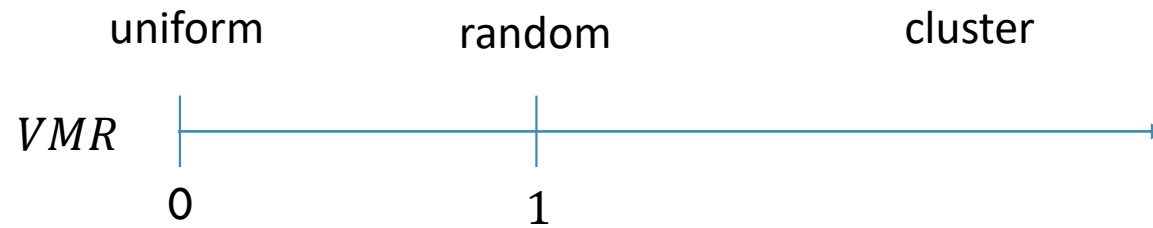
Step 1 - fishnet **GridTopology()**

Step 2 - calculate counts of points in each grid **poly.count()**

Step 3 - calculate **mean** and **variance** of counts

Step 4 - hypothesis testing: Variance-Mean Ratio Test (t test)

Step 5 - make a conclusion



## VMR Test

$$VMR = \frac{vairance}{mean} ; s.e. = \sqrt{\frac{2}{k-1}}$$

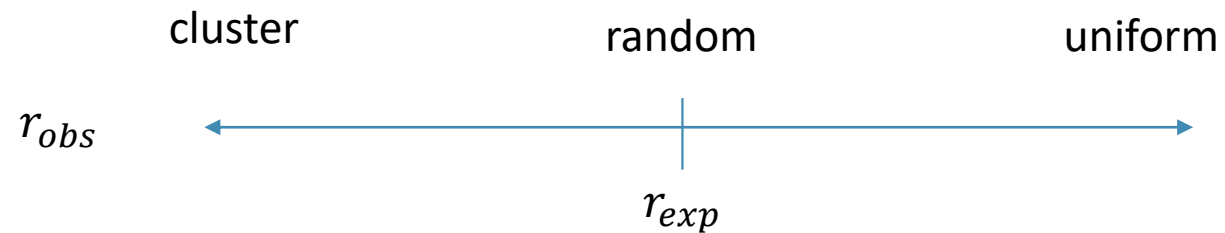
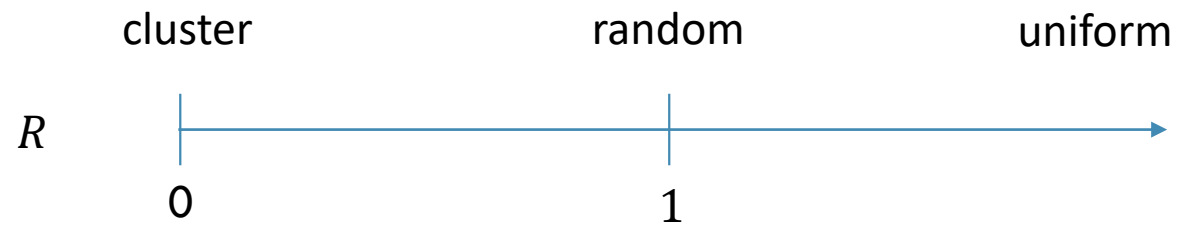
$$t = \frac{VMR - 1}{s.e.}, \quad df = k - 1$$

## Nearest Neighbor Analysis (NNA)

Step 1 - 每一個點，找最近的點的距離

Step 2 - 所有距離的平均，得到 $r_{obs}$

Step 3 - 觀察值與理論隨機值的比值： $R = r_{obs} / r_{exp}$



## K-order NNI

每一個點，找第k近的點的距離

## Nearest Neighbor Analysis (NNA)

### 顯著性檢定

#### 1. 理論隨機分布

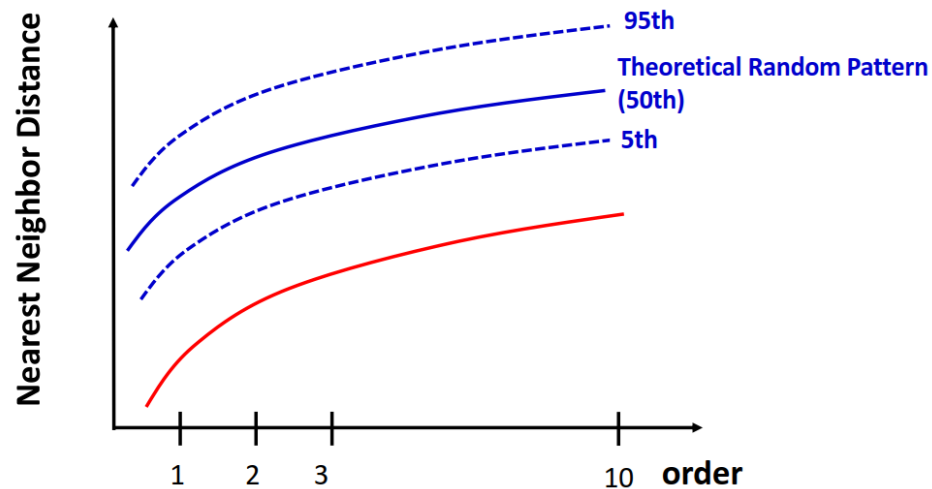
$$r_{exp} = \frac{0.5}{\sqrt{n/A}} ; s.e. = \frac{0.26136}{\sqrt{n^2/A}}$$

$$Z = \frac{r_{obs} - r_{exp}}{s.e.}$$

#### 2. Monte Carlo 顯著性檢定

模擬隨機分布 (i.e.1000次)

找出是否落在信賴包絡(confidence envelope)之中  
(i.e.1000次中, 排序前後25名的數值)





點轉換成ppp的格式

```
ppp(x.coor,y.coor,x.range,y.range)
```

最鄰近距離(NNA)

```
nndist(points.ppp,k=1) → mean
```

K order-NNI (i.e.最近第1~100點)

```
nndist(points.ppp,k=1:100)
```

```
apply(nndist(points.ppp,k=1:100),2,mean)
```

# 產生隨機點

Windows=**owin**(xrange=x.range, yrange=y.range)

RandomPoints=**rpoint**( **n** , win=Windows) → **n** : 個數

概念：

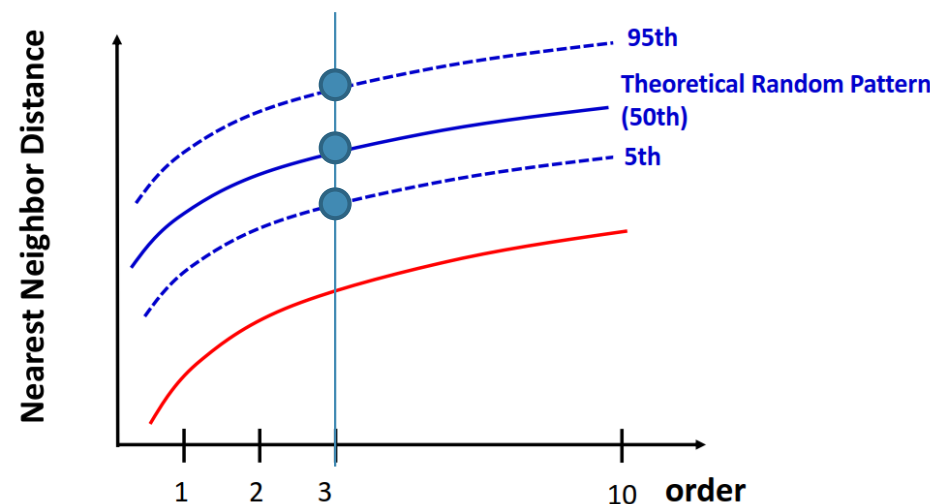
mean(nndist(points.ppp,k=**1**)) 會有1個數值

apply(nndist(points.ppp,k=**1:100**),2,mean)會有100個數值

模擬1000次：

1\*1000

100\*1000

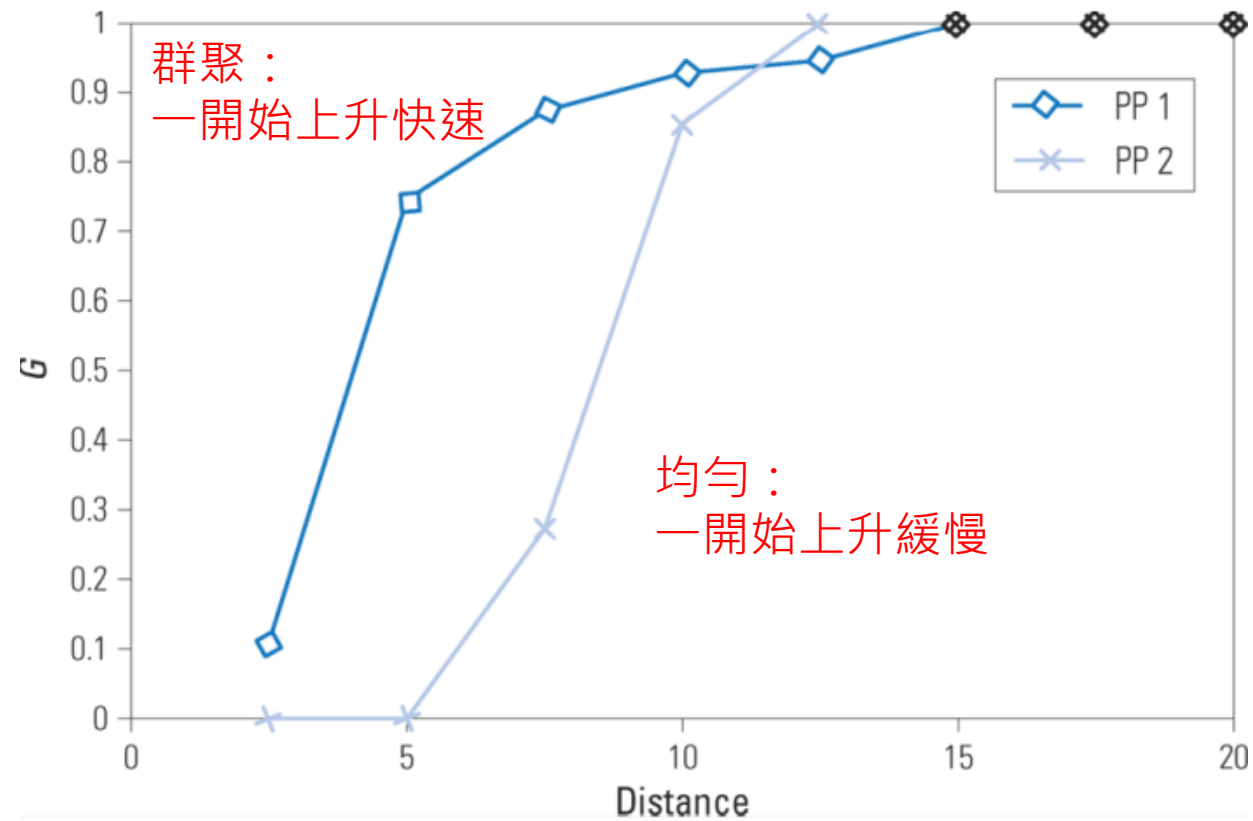
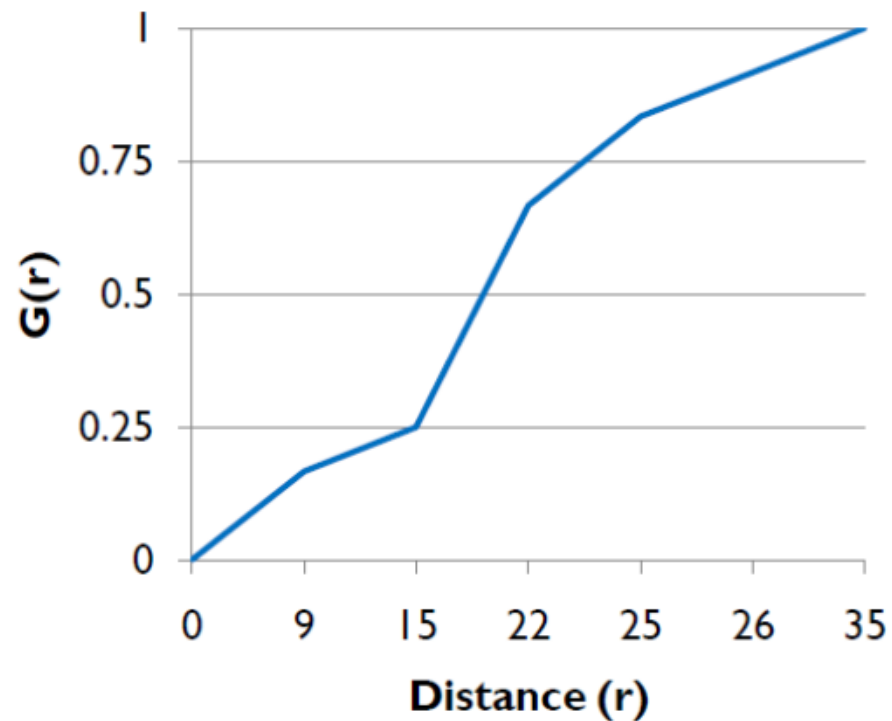


1000次找前後5%

排序函數：**sort()**

# G Function

「最近的點的距離」的累積頻率分布



## G Function 實作

```
nnd=nndist(School.ppp, k=1)
```

$G = \text{ecdf}(nnd)$  → 累積頻率分布

WindowPolygon= owin(poly=多邊形，用逆時針的點來表示)

ppp(x.coor,y.coor, window=WindowPolygon)

	x1	y1
[1,]	164481.9	2541476
[2,]	164497.4	2541324
[3,]	164390.8	2541256
[4,]	164288.9	2541191
[5,]	164241.2	2541161
[6,]	164214.4	2541046
[7,]	164223.8	2541030
[8,]	164223.2	2541027
[9,]	164214.4	2541018

