

## 108-2 空間分析 第二次期中考

考試時間：2020 年 5 月 18 日（一）下午 1:30~5:30

授課教師：溫在弘 | 課程助教：杜承軒、江偉銘、楊鎧綾

\* 作答提醒注意事項：

1. 本次考試採遠距考試，可自由上網查詢參考資料，但考生之間不得交流、共用交換資料以及代考，違規者將視同作弊，作弊者則考試成績將不計分。
2. 答案卷以 RMarkdown 格式輸出成 html 上傳，應於 5:30 pm 之前繳交（以 ceiba 上傳時間為準）；若上傳時間在 5:40pm 以後，則不予計分。
3. 請檢查各題目的作答要求，確認經 RMarkdown 輸出後，圖表、表格、答案等資訊有呈現在 html 中。所有試題皆以電腦作答，請隨時進行存檔。若因電腦當機或其他個人因素，因檔案未能及時存檔，導致無法準時交卷，請自行負責。

\* 成績計算說明：

[A 組]

- 一份考卷共 100 點，全班分配 4000 分。原始成績計算為「個人獲得點數  $\times$  4000  $\div$  A 組同學獲得總點數」，超過 100 分者以 100 分計。
- 原始成績會依第一週公告的分數調整後，再加上個人加分（課堂問題或加分題）。

[B 組]

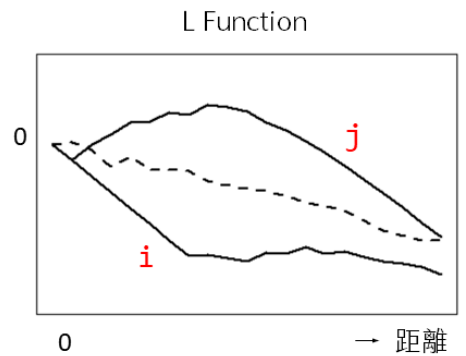
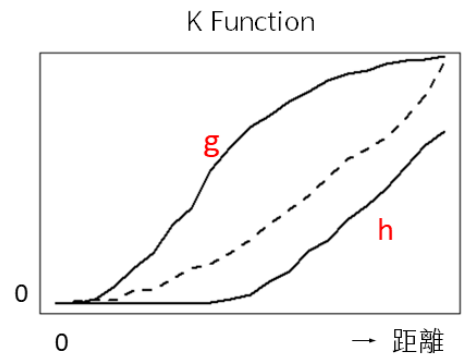
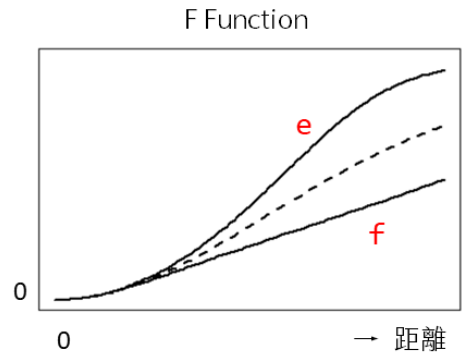
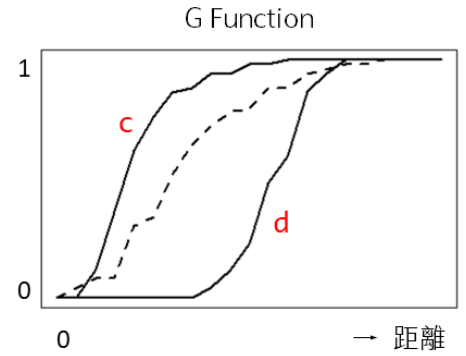
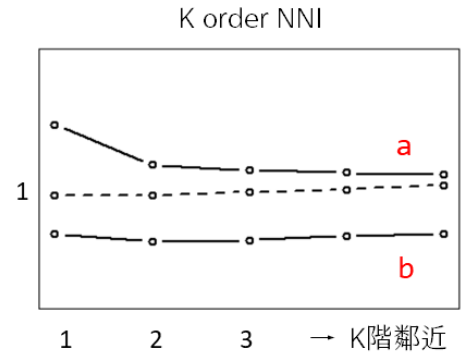
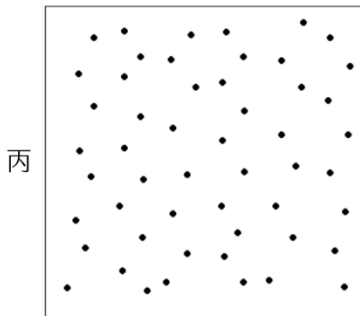
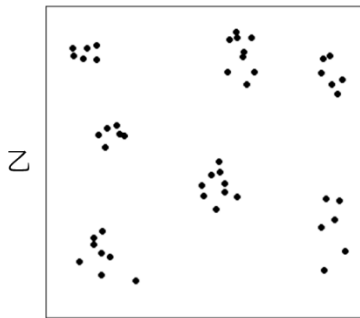
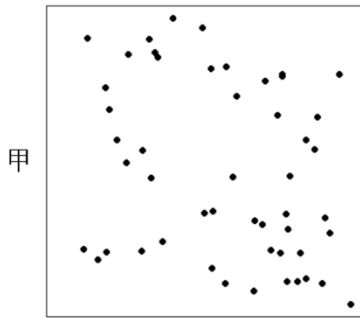
- 採計基礎題 50 點，個人獨立計分。

## 基礎題 (50 點)

### 一、圖表判讀 (10%)

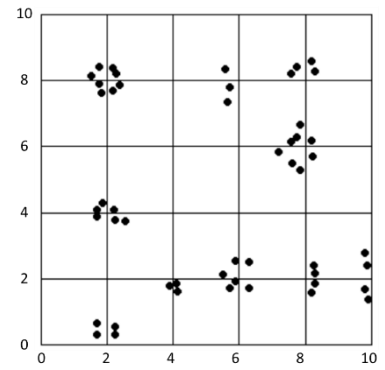
甲、乙、丙三個情境之點分布如下圖。

將甲、乙、丙進行點空間型態分析，得到右方五張圖表的結果，其中虛線對應到的皆是甲情境之點分布。請問 a~j 中，哪五條線代表的是乙情境之點分布？



## 二、病例點分布之分析 (30%)

**case.shp** 是某地本土個案感染者的住家點位，**grid.shp** 代表研究區的邊界範圍及研究者劃定的網格，兩者參考座標相同且單位為公里，如右圖。



在此研究中，研究者想知道本土感染者是不是隨機分布。設定虛無假設  $H_0$ ：病例點分布為隨機分布，對立假設  $H_a$ ：病例點分布不為隨機分布，顯著水準  $\alpha = 0.1$ 。請依照以下各小題的方法來回答問題。

### 1. 樣方分析 (quadrat analysis)

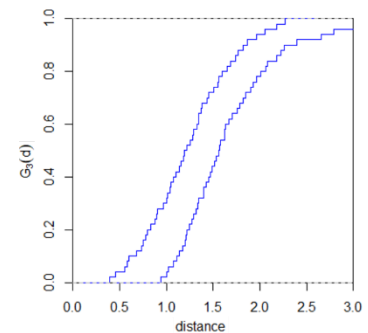
- (1) 計算 VMR (variance-mean ratio) 的數值。 (3%)
- (2) 透過理論公式進行 VMR 檢定，回答病例點是否呈隨機分布？ (3%)

### 2. 高階最鄰近距離分析 (K-order nearest neighbor distance)

- (1) 設定  $K = 1 \sim 5$ ，以  $K$  值為  $x$  軸，平均鄰近距離為  $y$  軸，繪製出前五階鄰近距離之關係圖。 (4%)
- (2) 實作蒙地卡羅方法，模擬 99 次隨機值的狀況，依照顯著水準的設定，在圖表上繪製信賴包絡曲線，並回答病例點是否呈隨機分布？ (6%)

### 3. 高階 G 函數

- (1) 定義  $G_3(d)$  為第三階鄰近距離的累積頻率分布。以距離為  $x$  軸， $G_3(d)$  為  $y$  軸，繪製出第三階鄰近 G 函數的曲線。 (4%)
- (2) 右圖為隨機模擬下  $\alpha = 0.1$  的信賴包絡曲線，透過蒙地卡羅檢定，觀察圖表，回答病例點是否呈隨機分布？ (2%)



### 4. Ripley's K 函數

定義  $L(d) = \sqrt{K(d) \div \pi} - d$ 。透過 spatstat 套件中 envelope 函數實作蒙地卡羅模擬，以距離  $d$  為  $x$  軸， $L(d)$  為  $y$  軸，繪製出  $L(d)$  函數曲線，以及模擬 99 次隨機值之信賴包絡 (需依照顯著水準的設定區間)，並回答病例點是否呈隨機分布？ (5%)

5. 綜合上述四個方法的檢定結果，你認為評估病例點是否呈隨機分布？如果不同方法之間的結果有差異，請加以敘述結果不一致的原因。 (3%)

### 三、程式碼判讀 (10%)

某次課程後，助教提供參考程式碼讓大家熟悉分析方法的運算過程，然而寄信時卻忘了附上方法的名稱。請觀察以下程式碼來回答問題。

```
1 # pts 格式為 SpatialPointsDataFrame
2 n = length(pts)
3 DIST = gDistance(pts,byid=T)
4 for(i in 1:n){ DIST[i,]= sort(DIST[i,]) }
5 d=DIST[,2]
6 x=seq(0,3000,100) #距離
7 y=c()
8 for(i in 1:length(x)){ y[i]=sum(d<x[i])/n }
9 plot(x,y,type='l')
```

1. 請由程式碼來判斷，上述是哪個點型態分析方法的實作？ (4%)
2. 第 5 行中 d 代表的意義為何？為何是設定 DIST 矩陣的第 2 欄？ (6%)

## 進階題 (50 點)

\*圖資：(座標系統：TWD97/TM2 121°)

**TPE.shp**：台北市範圍

**FASTFOOD.shp**：速食店點位

**SCHOOL.shp**：學校點位

■ Name：學校名稱

■ Type：公立／私立

■ index：該校學生平均腰圍身高比（該校肥胖指標）

※ 屬性資料皆是虛構，並非真實數值，僅作為學習評量目的

1. 公立與私立學校的平均腰圍身高比，是否有顯著差異（需列出假設檢定過程）？(5%)
2. 透過 KDE 方法，設定網格數量為 $100 \times 100$ ，核密度函數為 quartic，搜尋半徑為 1000 公尺，預計找出哪些學校周圍的速食店獨占性最高？即該地方附近的速食店多但學校數量少。

※ Hint: kernel2d 中的 poly 引數要放多邊形的座標

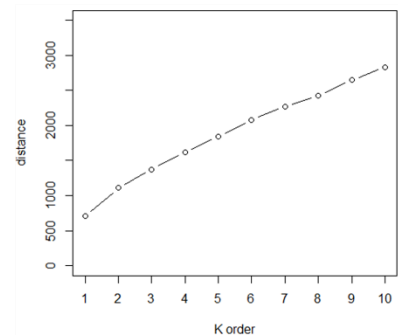
(1) 依上述指示畫出 Dual KDE 地圖。(6%)

(2) 列出速食店獨占性最高的三間學校之名稱（Hint: 學校所在網格之 KDE 數值）。(6%)

3. 找出每間學校到第 K 鄰近速食店的距離，並以第 K 鄰近為 x 軸，平均距離為 y 軸來製作圖表。設定 K 從第 1 鄰近至第 10 鄰近(K = 1~10)。

(1) 右方示意圖是「所有學校」到鄰近速食店之平均距離曲線，請用相同的概念，分別繪製「公立學校」（藍線）與「私立學校」（紅線）的曲線。(6%)

(2) 解釋上一小題兩條曲線的意涵。(2%)

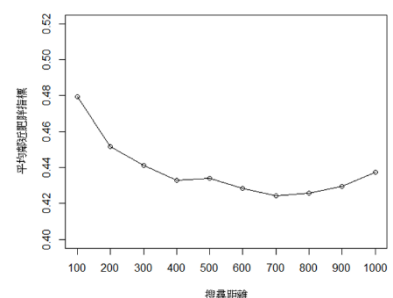
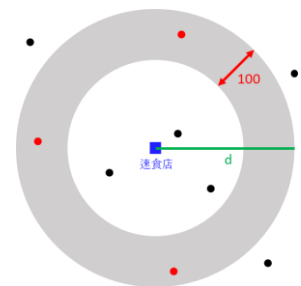


4. 透過速食店的搜尋範圍，來回答鄰近速食店與兒童肥胖的問題。定義搜尋距離為 d 時，搜尋範圍為 d - 100 至 d 的區域，如右圖的灰色部分。找出環形搜尋範圍的學校(紅色點位)，並計算所選學校肥胖指標之平均(即紅色點 index 數值之平均)，作為此速食店的鄰近肥胖指標。而平均鄰近肥胖指標是在搜尋距離為 d 時，所有速食店的鄰近肥胖指標之平均，但若該速食店之搜尋範圍中無學校，則忽略該店。

(1) 設定 d = 300，找出鄰近肥胖指標最高的速食店的 id 為何？鄰近肥胖指標為多少？所有速食店之平均鄰近肥胖指標為多少？(12%)

(2) 以 d = 100, 200, ..., 1000 來當作搜尋距離，以速食店的搜尋距離為 x 軸，平均鄰近肥胖指標為 y 軸來製作圖表，右方示意圖是搜尋「所有學校」，請用相同的概念，分別繪製「公立學校」（藍線）與「私立學校」（紅線）的曲線。(6%)

(3) 解釋上一小題兩條曲線的意涵。(2%)



5. 請根據前述的檢定結果與圖表，綜合評估速食店、公私立學校與學生肥胖的關連性。(5%)