5. التحليل المكاني للظواهر النقطية

ويشمل طريقتين من التحليل :-

اولا: التوزيع المكاني للنقاط:- وتتمثل بتحليل التوزيع المكاني للنقاط وتتضمن طرق عدة منها ، أ). طريقة الفرز بالمربعات quadrat counts analysis ب). طريقة تحليل الجار الاقرب Nearest neighbor methods ، جـ). الارتباط الذاتي autocorrelation وشمل معامل موران moran’s I ومعامل جيري **Geary's C** .

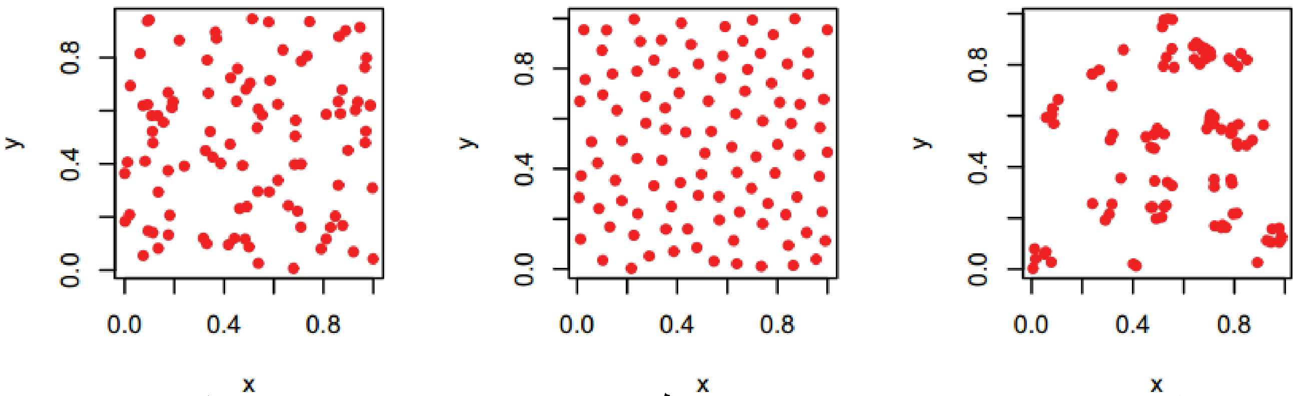
ثانيا: طريقة الحشر والاستكمال interpolation ، وتستخدم هذه الطريقة لتحديد التوزيع لقيمة البعد الثالث للنقاط بحسب توزيعها المكاني.، طريقة ثايسون، وزن معكوس المسافة ، كرجنك ، وكثافة النقاط وغيرها.

التوزيع المكاني للنقاط ويشمل:-

أ). طريقة الفرز بالمربعات:

وهنا يتم التركيز على النقاط ذات البعدين في التوزيع مثل مواقع الأشجار في الغابة، مركز الزلزال، والأنابيب البركانية، والينابيع، المتحجرات على الطبقات الصخرية، والرواسب المعدنية، وتفشي الأمراض وحوادث الجريمة .

والفكرة هي التمييز بين أنماط نقطة التي تميل كليا نحو عشوائية التوزيع المكاني، او الانتظامية او التوزيع العنقودي . Random, regular, and cluster distribution .



عشوائي　　　　منتظم　　　　عنقودي

-مؤشرات تحليل نمط توزيع النقاط Criteria of points pattern analysis

هناك مؤشرات يجب توفرها في البيانات كي يتسنى لنا فحص نمط التوزيع النقطي لها وهي:-

1. يجب أن يتم تعيين البيانات المكانية على مستوٍ، ويحتاج كلا الإحداثيات لخطوط الطول والعرض للنقطة.

2. يجب تحديد منطقة الدراسة وتحديد مسبق لنوع التحليل.
3. لا ينبغي أن تكون البيانات عينات نقاط مختارة، بل مجموعة كاملة من البيانات المراد تحليلها.
4. ينبغي أن تكون هناك علاقة بين العوارض في منطقة الدراسة والاحداث في النمط.
5. يجب أن تكون النقاط حوادث حقيقية ولها الإحداثيات المكانية الحقيقية.

1. طريقة الفرز الرباعي quadrat counts analysis

وهذه طريقة تعتمد على كثافة النقاط ، وتتم بحساب عدد النقاط او تكراراتها ضمن شبكة من المربعات .

- تحليل الرباعي يقيم توزيع النقاط من خلال فحص كيفة تغير كثافتها على الفضاء.
- ثم تتم مقارنة كثافة قياس تحليل المربعات مع كثافة نمط عشوائي مشيد بطريقة نظرية لمعرفة ما إذا كان توزيع النقاط هل هو أكثر عنقودي ام متشتت بعيدا عن النمط العشوائي.
- انشاء شبكة مربعات منتظمة تقع في بعضها عدد من النقاط .
- ويشار إلى المربعات باسم الأرباع، التي هي أساس نمذجة العينات في التحليلات الاحصائية المكانية.
- في نمط النقاط العنقودي للغاية، كل أو معظم النقاط تقع داخل واحد أو عدد قليل من المربعات فقط.
- في نمط الغاية في التشتت فان كل المربعات تحتوي على عدد مماثل من النقاط.

ويتم حسابها وفق الاتي:-

1. رسم شبكة من المربعات تغطي منطقة الدراسة
2. حساب عدد النقاط في كل مربع .
3. اختيار حجم المربع لرسم شبكة المربعات مهم جدا في دقة الحساب ونتيجة التحليل اذ يجب ان يكون منسابا للتحليل .
4. يتم حساب مساحة المربع ومنها يحسب طول ضلع المربع باستخدام المعادلة الاتية:-

$$areaofsquare = 2\frac{Area}{n}$$

$$widthofsquare = \sqrt{areaofsquare}$$

where ***n*** is the number of points in the sample size.

5. استخدام مؤشر نسبة التغاير الى المعدل لتحديد مدى الابتعاد عن كل من التوزيع المنتظم والعشوائي والعنقودي، وكما ياتي :-

$$Mean = \frac{No.\ of\ \ pts.in\,the\,region}{No.of\ \ quadrants}$$

$$Variance = \frac{\sum x_i^{\ 2} - \frac{(\sum x_i)^2}{n}}{n-1}$$

x_i يمثل تكرار او عدد النقاط في كل مربع

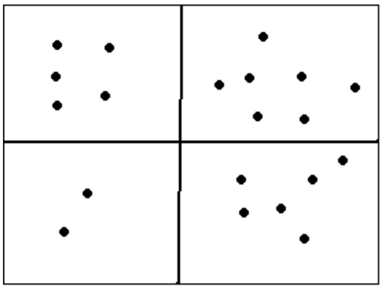
The Variance to Mean Ratio or VTMR is calculated as:

$$VTMR = \frac{Variance}{Mean}$$

بعدها يتم مقارنة قيمة VTMR مع الواحد الصحيح (1) ليتم بعدها تصنيف توزيع النقاط على انه اما منتظم او عنقودي او عشوائي .

- اذا كانت قيمة VTMR اكبر من واحد (1) فان التوزيع عنقودي clustered .
- اذا كانت قيمته اقل من واحد (1) فان التوزيع متشتت dispersed
- اما اذا يساوي واحد (1) فان التوزيع عشوائي .

مثال:-



$$Mean = \frac{No.\ of\ \ pts.in\,the\,region}{No.of\ \ quadrants} = \frac{20}{4} = 5$$

x_i يمثل تكرار او عدد النقاط في كل مربع

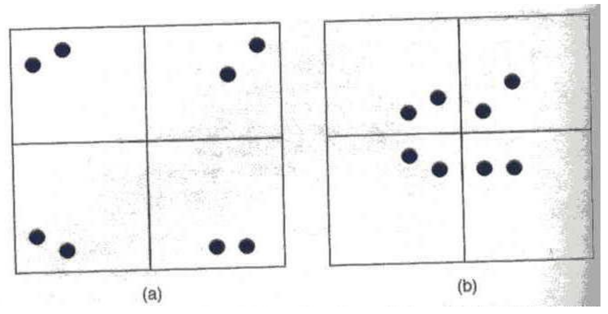
$$Variance = \frac{\sum x_i^{\ 2} - \frac{(\sum x_i)^2}{n}}{n-1} = \frac{2^2 + 5^2 + 6^2 + 7^2 - \frac{(20)^2}{4}}{4-1} = 4.5$$

$$VTMR = \frac{Variance}{Mean} = \frac{4.5}{5} = 0.9$$

هذا يعني ان نمط توزيع النقاط هو متشتت

2. تحليل الجار الاقرب Nearest Neighbor Analysis

تحليل الفرز الرباعي مفيد في مقارنة بين توزيع النقاط الملاحظة مع توزيع عشوائي نظري. ومع ذلك، فإنه يحتوي على بعض القيود. حيث انه يعطي معلومات عن النقاط في كل مربع، ولكن ليس فيه اشارة للنقاط في المربعات المجاورة في التحليل، لذلك فهو محدود الامكانيات وكما مبين في الشكل الاتي .



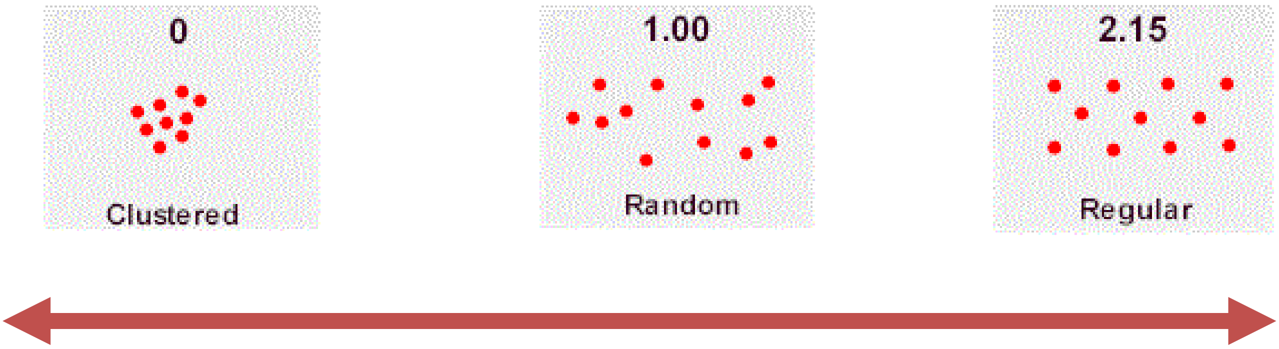
من الملاحظ ان الشكلين يختلفان في نمط توزيع النقاط ، ولكن بطريقة التحليل الرباعي النتيجة هي واحدة ، أي في كل مربع توجد نقطتين

لذلك فان تحليل الجار الاقرب يوفر حلا لهذا الاشكال .

- يتطلب تحليل الجار الاقرب وجود 30 نقطة فاكثر كي يكون دقيقا .
- البيانات الاحصائية لتحليل الجار الاقرب يتم اشتقاقها من معدل المسافة بين النقاط واقرب جار لها .
- الجار الاقرب بالرتبة الثانية يعتمد على حساب معدل المسافة بين والنقاط واقرب ثاني جار لها . وهكذا الحال لتحليل الجار الاقرب بالرتب العالية .
- بالرغم من ان كلا التحليلين (الجار الاقرب ، وتحليل الرباعي) يعملان على تحليل توزيع النقاط الا انهما يعملان وفق مفهومين مكانيين مختلفين.
- *التحليل الرباعي يفحص توزيع النقاط وفق مفهوم حساب عدد النقاط داخل مساحة المربع الواحد . اما تحليل الجار الاقرب فانه يعتمد مبدأ المساحة الى النقاط .*

تحليل الجار الاقرب ويرمز له Rn هو نوع اخر من تحليل نمط توزيع النقاط ويقوم بفحصها هل نمط توزيعها هو عنقودي او عشوائي او منتظم .

- *التوزيع العنقودي Cluster pattern يكون عندما كل النقاط تتوزع بالقرب من نقطة واحدة بشكل عنقودي أي ان قيمة* ***Rn = 0***
- *التوزيع العشوائي Random pattern : ويحصل عندما لايوجد نمط معين لتوزيع النقاط أي ان قيمة* ***Rn = 1*** *وعادة توزيع النقاط في الطبيعة يكون عشوائيا مع ميل نحو العنقدة او الانتظامية .*
- *النمط المنتظم regular pattern وفيه يكون توزيع النقاط منتظم تماما وقيمة* ***Rn =2.15*** *والنقاط تتوزع بمسافات متساوية.*



Clustered **Random** **Regular**

(nucleated)

tendency towards tendency towards **(uniform)**

$$R_n = 2 \times \bar{d}_{obs} \times \sqrt{\frac{n}{A}}$$

Where

Rn = وصف التوزيع بحسب الجار الاقرب

$\bar{d}_{obs}$ = معدل المسافة بين النقاط (بحسب وحدة القياس)

A = مساحة منطقة الدراسة (مربع وحدة القياس)

n = عدد النقاط الكلي

$$\bar{d}_{obs} = \frac{\sum di}{n}$$

di= المسافة بين كل نقطة واقرب نقطة ومجاورة لها

مثال:-
منطقة مساحتها 12 كم * 15كم ومجموع المسافات بين النقاط بحسب جوارها هو 51.7 كم وعدد النقاط هو 30 نقطة. جد قيمة Rn .
الحل :-

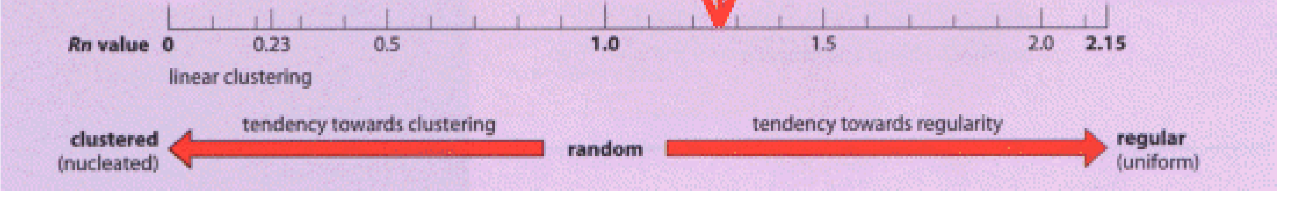
$$\bar{\boldsymbol{d}}_{\boldsymbol{obs}} = \frac{\sum di}{n} = 51.7/30 = 1.72 \text{ km}$$

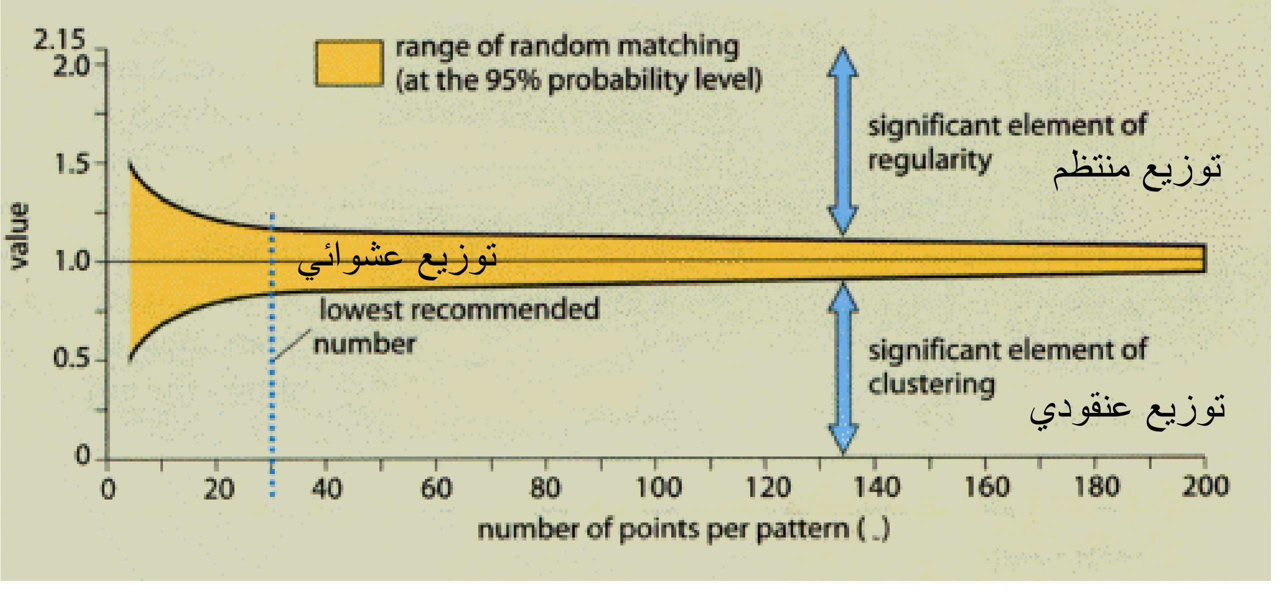
مساحة المنطقة 12× 15 = 180كم2

$$R_n = 2 \times \bar{d}_{obs} \times \sqrt{\frac{n}{A}}$$

Rn = 2 x 1.72 √ 30/ 180 = 3.44 √ 0.17 = 3.44 x 0.41= 1.41

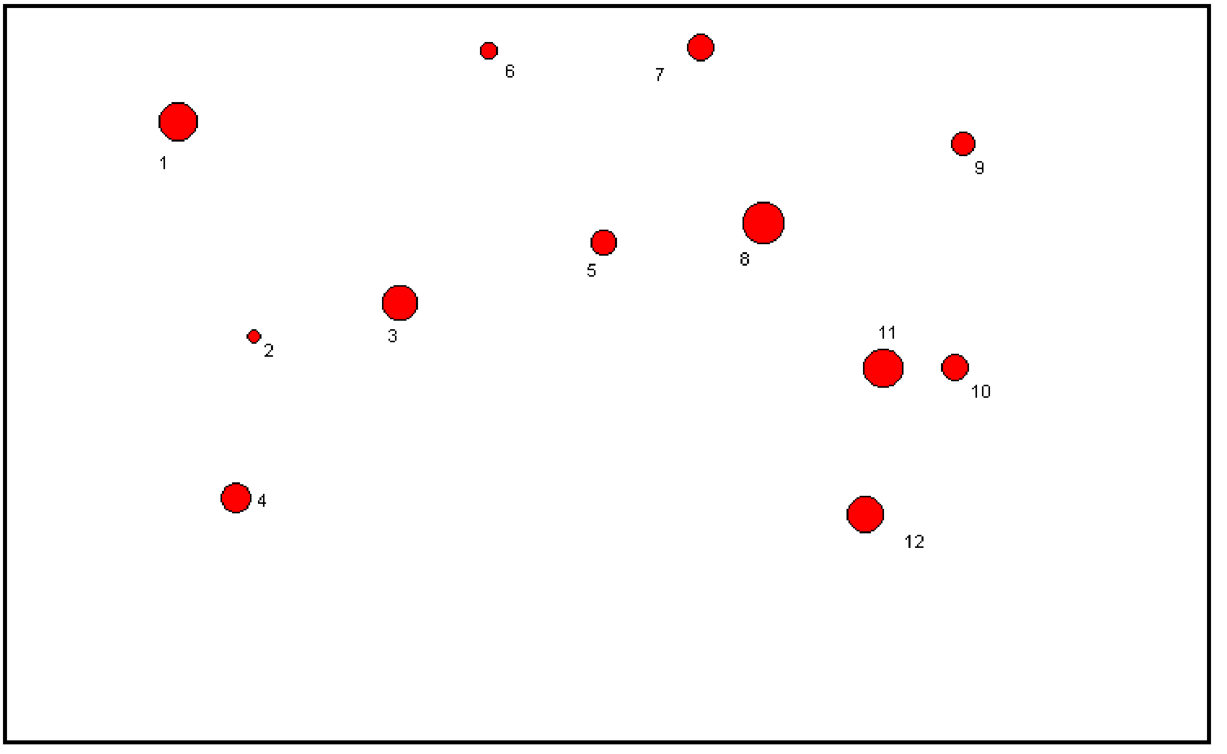
1.41





يبين المخطط ان المنطقة المظللة تمثل مستوى احتمالية 95% توزيع عشوائي ، وان قيمة Rn يجب ان تقع خارج المنطقة المظللة كي يكون التوزيع اتجاهه نحو التوزيع العنقودي او التوزيع المنتظم .

مثال 2:- جد نمط التوزيع للنقاط الاتية باستخدام طريقة الجار الاقرب.



$$R_n = 2 \times \bar{d}_{obs} \times \sqrt{\frac{n}{A}}$$

$$\bar{d}_{obs} = \frac{\sum di}{n}$$

where *di*= المسافة بين النقطة وجارتها

النقطة	وجارها الاقرب	المسافة (كم)
1	2	13.0
2	3	9.0
3	2	9.0
4	2	9.5
5	8	9.0
6	7	12.5
7	6	12.5
8	5	9.0
9	8	12.5
10	11	4.0
11	10	4.0
12	11	8.5

Solving nearest neighbor by matrix method:
حل الجار الاقرب بطريقة المصفوفة

Nearest neighbor index (R ratio)

$$R = \frac{\bar{d}_{obs}}{\bar{d}_{ran}}$$

R = degree of clustering

$$\bar{d}_{obs} = \frac{\sum d_i}{N}$$

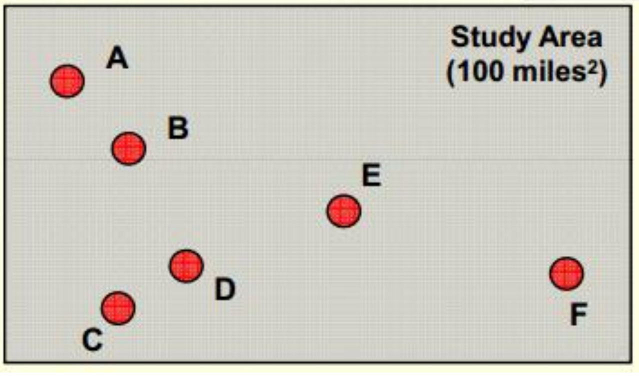
d_i = distance to nearest neighbor of point i, and

N = # points

$$\bar{d}_{ran} = \frac{1}{2\sqrt{\rho}}$$

ρ = density of points per unit area

$$\rho = \frac{n}{A}$$



Step 2:

Distance Matrix (Miles)

	A	B	C	D	E	F
A	0	2	8	6	7	8
B	2	0	6	4	5	6
C	8	6	0	3	7	6
D	6	4	3	0	4	6
E	7	5	7	4	0	5
F	8	6	6	6	5	0

Lowest values in each column (nearest neighbors)

Step1:

Distance Matrix (Miles)

	A	B	C	D	E	F
A	0	2	8	6	7	8
B	2	0	6	4	5	6
C	8	6	0	3	7	6
D	6	4	3	0	4	6
E	7	5	7	4	0	5
F	8	6	6	6	5	0

Step 3:

Distance Matrix (Miles)

	A	B	C	D	E	F
A	0	2	8	6	7	8
B	2	0	6	4	5	6
C	8	6	0	3	7	6
D	6	4	3	0	4	6
E	7	5	7	4	0	5
F	8	6	6	6	5	0

$\bar{d}_{obs}$ = (2+2+3+3+4+5)/6 = 3.17 miles

$\bar{d}_{ran}$ = 1/(2 $\sqrt{(6\ points\ /\ 100\ miles^2)}$) = 2.04 miles

So, finish the R calculation

$$R=\frac{\bar{d}_{obs}}{\bar{d}_{ran}}=\frac{3.17}{2.04}=1.55$$

Interpretation:

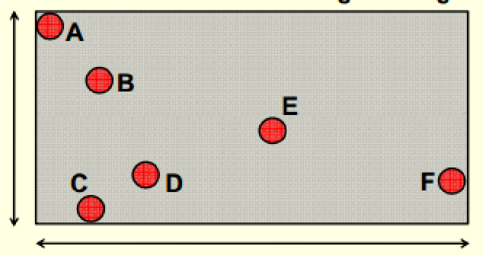
R=1.0: random

R=0.0: clustered

R=2.1491: dispersed (uniform)

maximum possible

"Enclosing Rectangle"



Minimum height and width of rectangle that encloses all points

- **Key Question: Significance**
  - Test Statistic ("Geary's C"):

$$C=\frac{(\bar{d}_{obs}-\bar{d}_{ran})}{SE_{\bar{d}}}$$

Standard error of the NN distance

- **Key Question: Significance**
  - Standard error calculation

$$SE_{\bar{d}}=\frac{0.26136}{\sqrt{N\times\rho}}$$

N = # points

ρ = density of points per unit area

- Caution: use the same area in every period for a time-based study (examining changes in a study area) – difficult to compare R values when the study area changes in some way

- **Key Question: Significance**
  - Is the pattern significantly different from random?
  - H_0: pattern is random (always have same H_0)
- **Options come in on your H_1 (choose one)**
  - H_1: pattern is not random (two-tailed test)
  - H_1: pattern is not random and is clustered (one-tailed test)
  - H_1: pattern is not random and is dispersed (one-tailed test)

- **Key Question: Significance**
  - Test Statistic ("Geary's C"):

$$C=\frac{(\bar{d}_{obs}-\bar{d}_{ran})}{SE_{\bar{d}}}$$

Standard error of the NN distance

- **Key Question: Significance**
  - Doing the calculation with the "6 dots on a map" example and the associated values calculated earlier

$$SE_{\bar{d}}=\frac{0.26136}{\sqrt{N\times\rho}}=\frac{0.26136}{\sqrt{6\times 0.06}}=0.43$$

  - Therefore, the value of the test statistic C is

$$C=\frac{\bar{d}_{obs}-\bar{d}_{ran}}{SE_{\bar{d}}}=\frac{3.17-2.04}{0.43}=2.63$$

- Compare the calculated C value with the critical value for the statistic (one tailed test, 0.05 level)

C10 Critical Values of a Standard Normal Deviate z

	Significance level (one-tailed)				
	0.1	0.05	0.01	0.005	0.001
z	1.282	1.645	2.326	2.576	3.090
−z	−1.282	−1.645	−2.326	−2.576	−3.090
	Significance level (two-tailed)				
	0.1	0.05	0.01	0.005	0.001
z	1.645	1.960	2.576	2.813	3.291
−z	−1.645	−1.960	−2.576	−2.813	−3.291

- From the table, C_{Crit}=1.645 (remember C_{Calc}=2.63)
- So, reject H_0 ($C_{crit}<C_{Calc}$): the pattern is significantly uniform and not random