To evaluate the geographic distribution by residence at the time of illness, cases from 1978 to 1981 within Miami-Dade County, a ciguatera endemic region, were analyzed (Figure 1). Of the 304 index cases, 189 occurred in Miami-Dade County, with 102 (60.4% of Miami-Dade County cases) of these cases occurring during the specified time period. A nearest-neighbor analysis was performed in an attempt to show a random distribution of cases in the county. However, despite various attempts to adjust for population density and lack of habitability (e.g., airports, Everglades, and ocean areas), the R-value was 0.10, indicating a strong clustering pattern. Nevertheless, the clustering pattern closely followed densely populated roadways that pass through highly varied neighborhoods.

Nearest-Neighbor Analysis

- Unlike quadrat analysis uses distances between points as its basis
- The mean of the distance observed between each point and its nearest neighbor is compared with the expected mean distance that would occur if the distribution were random
- Also needs a reference area
Advantages of Nearest Neighbor over Quadrat Analysis

- No quadrat size problem to be concerned with
- Takes distance into account
- Problems
  - Related to the entire boundary size
  - Must consider how to measure the boundary
    - Arbitrary or some natural boundary
    - May not consider a possible adjacent boundary
NNS problems

- NNS = 0.1
- NNS = 1.0
- NNS = 2.0
- NNS = 2.14

Does A relate geographically to B

Spatial Correspondence
- Coefficient of areal correspondence (Set theory, intersection / union
- Chi-square
- Yule’s Q

Clustering

Measurement
Compare over time, scale, theme, etc.

Overlay

Text Q example

Example Test of Spatial Pattern

• Is there a relationship between the distribution of rainfall and the wheat yield in the area shown?
• NULL HYPOTHESIS: There is no relationship
• ALTERNATIVE HYPOTHESIS: There is a relationship
**Chi-square**

- Make assumption that there is no relation between maps A and B
- Compute statistics that allow the assumption to be rejected
- Chi-square is the sum of the \((\text{Observed value} - \text{Expected value})^2/\text{Expected value}\)
- Can check value against table for actual likelihoods

---

### Calculating Chi-Squared: text p193

**Observed frequencies**

<table>
<thead>
<tr>
<th>Right: Wheat Yield</th>
<th>High</th>
<th>Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>8</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Low</td>
<td>5</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>13</td>
<td>15</td>
<td>28</td>
</tr>
</tbody>
</table>

---

### Calculating Chi-Squared: text p193

**Expected frequencies**

<table>
<thead>
<tr>
<th>Right: Wheat Yield</th>
<th>High</th>
<th>Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Low</td>
<td>8</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>13</td>
<td>15</td>
<td>28</td>
</tr>
</tbody>
</table>

*e.g. High yield is 10/28 cells = 35.7% times the total high yield of 13 = 5*
Calculating Chi-Squared: text p193

Observed and Expected differences

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8-5=3</td>
<td>2-5=3</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>[9/5]</td>
<td>[9/5]</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>5-8=-3</td>
<td>13-10=3</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>[9/8]</td>
<td>[9/10]</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>15</td>
<td>28</td>
</tr>
</tbody>
</table>

**Chi-squared**

\[
\text{Chi-square} = \sum \frac{(O-E)^2}{E}
\]

For the example = \(\frac{9}{5} + \frac{9}{5} + \frac{9}{8} + \frac{9}{10}\)
\[1.8 + 1.8 + 1.125 + 0.9 = 5.625\]

This value is then compared to a table of chi-squared to see if the value allows us to reject the null hypothesis that the observed values are not those expected based on proportions.

**Chi-squared tables**

Two by two table has four values so three degrees of freedom

Chi-squared of zero is no relationship.
Higher the value the stronger the relationship.

**Conclusion**

- Using Chi-squared it is not possible to reject the NULL hypothesis that there is no relationship between wheat yield and high precipitation
- Test statistic fails, but only just
- Use another method?
Yule’s Q

- Divide world into high/low (2 classes)
- Overlay two maps gives four classes
- Count quadrats in the four classes in a 2 x 2 table (with cells a,b,c,d) (i.e. Observed only)
- \( Q = \frac{(ad - bc)}{(ad + bc)} \)
- Value lies between -1 and +1
- -1 is perfect inverse relationship, +1 is perfect positive

Calculating Yule’s Q : text p193

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>8 (a)</td>
<td>2 (b)</td>
<td>10</td>
</tr>
<tr>
<td>Low</td>
<td>5 (c)</td>
<td>13 (d)</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>15</td>
<td>28</td>
</tr>
</tbody>
</table>

Calculating Q

\( Q = \frac{(ad - bc)}{(ad + bc)} \)

\( \frac{(8 \times 13) - (2 \times 5)}{(8 \times 13) + (2 \times 5)} = \frac{94}{114} = 0.82 \)

Close to +1, so can conclude that there is a positive relationship

Testing spatial relationships

- Is there a relationship between geographical location and the price of gas?
- Are apartment rents less as distance from the campus increase?
- Are grocery store prices higher in poorer areas?
- Are the increased cancer death rates in a district caused by water contamination?
- Is there a relationship between hydrocarbon emissions and decreased upper atmosphere ozone in the polar regions?
Summary

- Distributions can be quantified, using NNS or other means
- Maps can be compared using Chi-squared, Yule’s Q etc.
- Allows cartometry of higher order structures on maps: shape, distribution, arrangement and pattern