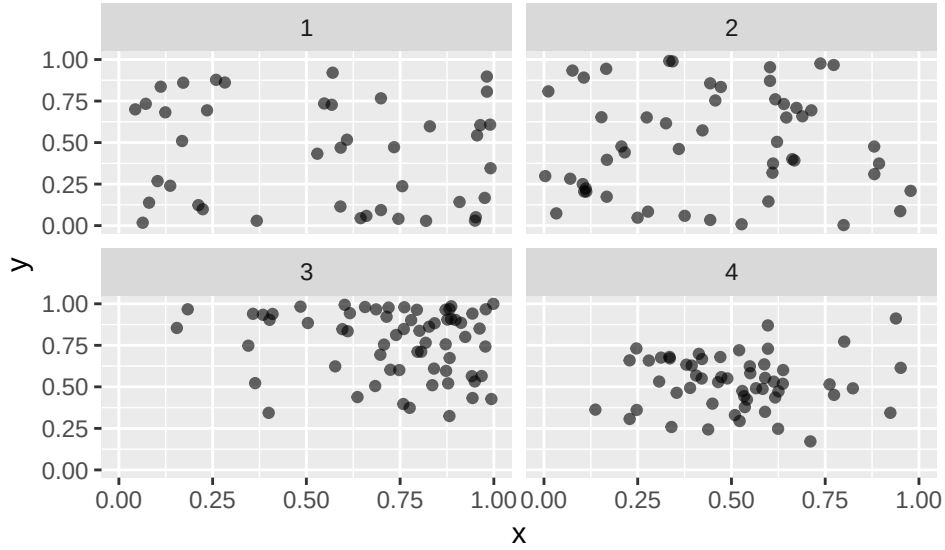
STAT 534 - Lecture 22: Key

Hypothesis Tests for CSR

- With a spatial point process, *a single realization from the underlying point process is observed.*

- Hypothesis tests and Monte Carlo simulation procedures allow testing for departures from Complete Spatial Randomness (CSR).

- Devise and implement a Monte Carlo-based test for CSR. First detail your test statistic and implementation, then assess the results with the four data sets from Lecture 21.



G and F Functions

- One way to describe a spatial point process, is to consider the probability of being a certain distance from a point or similarly, the number of points expected in a distance from a point.

- *Define* $Num(\boldsymbol{s}, d; \boldsymbol{S})$ *as the number of points from* $\boldsymbol{S}$ *in a circle of diameter* d *from point* $\boldsymbol{s}$*.*

- Then we can compute the probability that more than one point exists with a distance d of an observed point.

$$E_{\boldsymbol{S}}\left(\sum_{\boldsymbol{s_i}\in\boldsymbol{S}} 1(Num(\boldsymbol{s}_i), d; \boldsymbol{S}) > 0\right) = \lambda|\boldsymbol{D}|P(Num_d(\boldsymbol{s}, d; \boldsymbol{S}) > 0)$$

- The term $P(Num_d(\boldsymbol{s}, d; \boldsymbol{S}) > 0$ is referred to as $G(d)$. $G(d)$ increases in d.

- *The* $G(d)$ *function is simiilar to a CDF of the* ***nearest neighbor*** *distance*

- A similar statistic is the $F(d)$ function. Whereas $G(d)$ is centered at the observed $\boldsymbol{s}_i$, $F(d)$ is defined at any arbitrary point. *Hence this is a CDF for empty space.*

- **QUESTION:** Assume a CSR Poisson process with constant intensity λ, calculate $F(d)$ or $G(d)$ (they are the same).

- Discuss how to create an empirical estimate of $\hat{G}(d)$, given a realization of a point process.

- With bounded area, edge correction procedures are necessary. *Hence, consider*

$$\hat{G}(d) = \frac{\sum_i 1(d_i \leq d < b_i)}{\sum_i 1(d < b_i)},$$

where b_i is the distance from $\boldsymbol{s}_i$ to the edge and d_i is the distance to the nearest neigbor of $\boldsymbol{s}_i$.

- The empirical estimates of G or H can be compared with G or F using a QQ-plot.

- **Discuss:** What would be the implications of shorter tails or longer tails than expected under CSR? *shorter tail = clustering/attraction, longer tail = inhibition/repulsion*

- *Describe a natural process that might cluster and another than might repel*