**Spatial**

Ripley’s K, Besag’s L, and Marcon’s M quantify the of the degree of spatial clustering for a particular cell type. These quantities assume that the point process of interest follows complete spatial randomness (CSR) that is that the location of the cells do not form clusters nor do they occur in a regular pattern. Each measure is computed by counting the number of neighboring cells for each cell, where two cells are said to be neighboring if they are within a specified distance, r as seen in the x-axis of ‘Plot of Spatial Clustering Estimator’, of each other.

Border corrections are needed for these methods to account for the fact that these cell types are assumed occur outside of the TMA at the same rate. The edge corrections inflate the number of neighboring cells based on their proximity to the border and other cells. Two popular methods border corrections are isotropic and translational. Translation border correction translates the region of interest based on the distance between two points and then measures the area of the intersection of the original region and translated region. An isotropic border correction weights each pair of points based on how much of the circumference of a circle centered around one point and going through the other is outside of the region of interest.

Ripley’s K is computed by

where n is the number of cells, A is the area of the TMA, d(xi,dxj) is the distance between the ith and jth cell, and wij is weighting factor from the border correction. The expected value of K under CSR is Both Besag’s L and Marcon’s M are slight modifications to Ripley’s K and both more easily interpretable than Ripley’s K.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Formula | Expected Value | Formula Displayed in Plot | Interpretation |
| Besag’s L |  | r |  | If value is larger than 0, than there is evidence of spatial clustering. |
| Marcon’s M |  | 1 |  | If value is larger than 1, than there is evidence of spatial clustering. Additionally, if is 1.50 at some chosen radius means you’re observing 50% more clustering than the Poisson distribution would predict. |

**Summary**

The order of the samples (columns) and markers (rows) of the heatmap are determined by hierarchical clustering which assigns clusters based on ‘complete’ dissimilarity. Complete dissimilarity refers to the distance of two clusters being determined by the largest distance between elements in either cluster. The dataset will be sorted by the selected clinical variable to allow for a heatmap where the samples with the same class of a clinical variable are grouped together or appear in increasing order for continuous variables (such as age).