

Assignment 8

Colin Eberl-Coe

5 April 2016

1. Intro to spatstat.

- For each of $\lambda = 10, 20, 100$, generate five realizations of CSR on the unit square

10



10



10



10



10



20



20



20



20



20



100



100



100



100



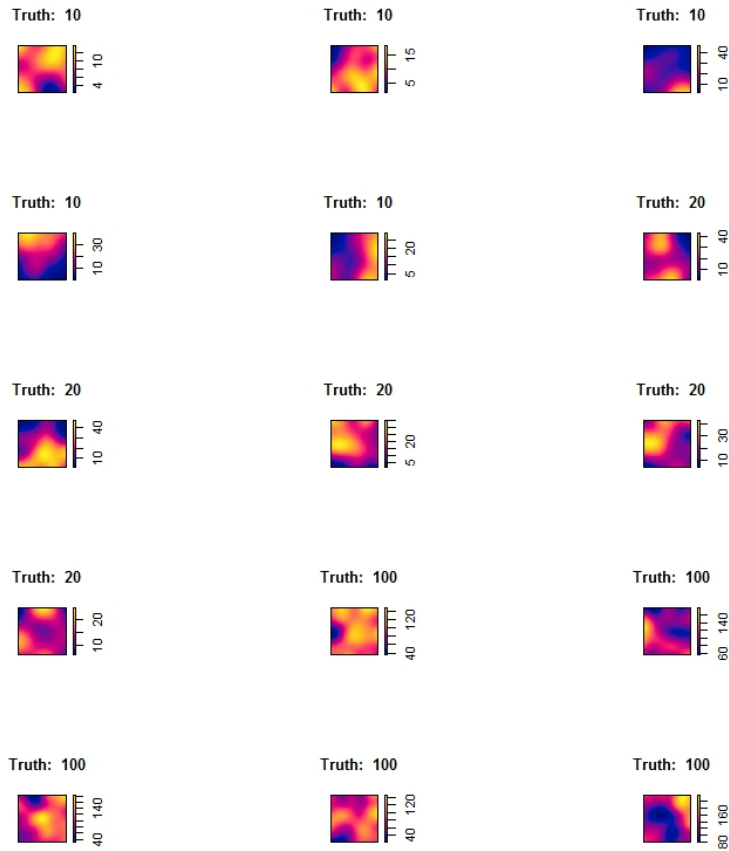
100



- Compute and report $\hat{\lambda}(\vec{s})_i$, b_i for $i = 1, 2, 3$; the values of lambda used above.

| ## | lam.true | lam.hat | b |
|-------|----------|---------|-----------|
| ## 1 | 10 | 9 | 0.2112789 |
| ## 2 | 10 | 11 | 0.1839476 |
| ## 3 | 10 | 15 | 0.1881779 |
| ## 4 | 10 | 6 | 0.2150985 |
| ## 5 | 10 | 12 | 0.1681937 |
| ## 6 | 20 | 19 | 0.1610956 |
| ## 7 | 20 | 24 | 0.1474580 |
| ## 8 | 20 | 15 | 0.1715253 |
| ## 9 | 20 | 21 | 0.1636150 |
| ## 10 | 20 | 20 | 0.1708288 |
| ## 11 | 100 | 123 | 0.1272803 |
| ## 12 | 100 | 127 | 0.1300260 |
| ## 13 | 100 | 107 | 0.1314484 |
| ## 14 | 100 | 104 | 0.1301507 |
| ## 15 | 100 | 83 | 0.1463259 |

- Plot the nonparametric estimates of λ using b found above.

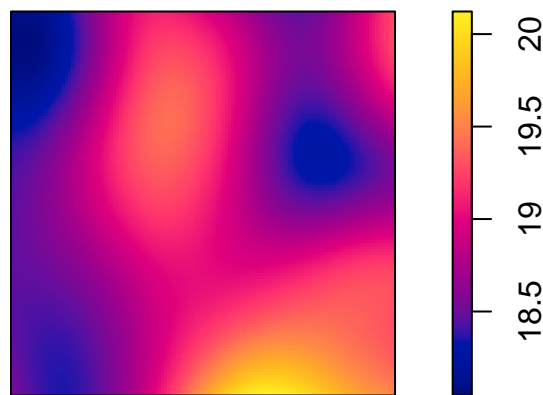


- How do the sets of five estimates of $\lambda(\vec{s})$ compare to the known constant values of λ ?

The estimate of λ is n/D , the number of points generated divided by the area of the domain, and is close to the known value of λ used in the simulation, as the above table demonstrates. However, looking at plots of the spatially-varying intensity returned by R, even though the intensity is known to be constant, the five realizations are different enough that interpreting a single realization could very easily lead to incorrect inference.

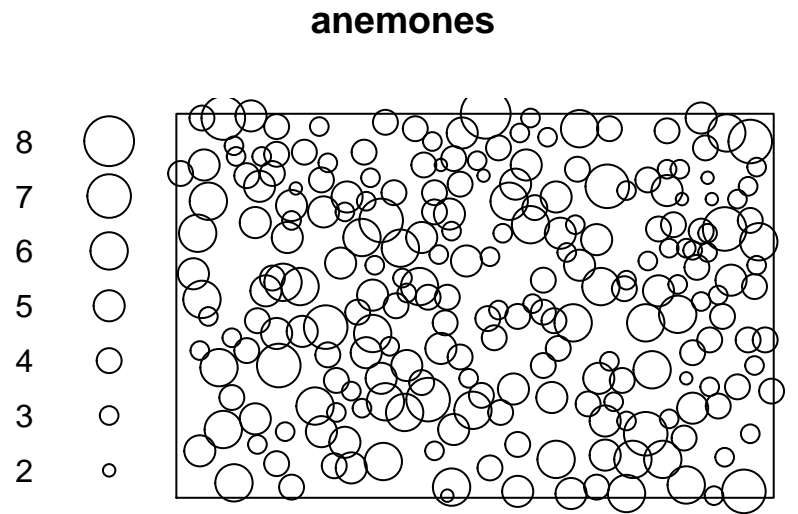
- Simulate 100 realizations of HPP and average their nonparametrically-estimated intensities. Plot this and compare.

Average of 100 simulations of $\hat{\lambda}$ per pixel

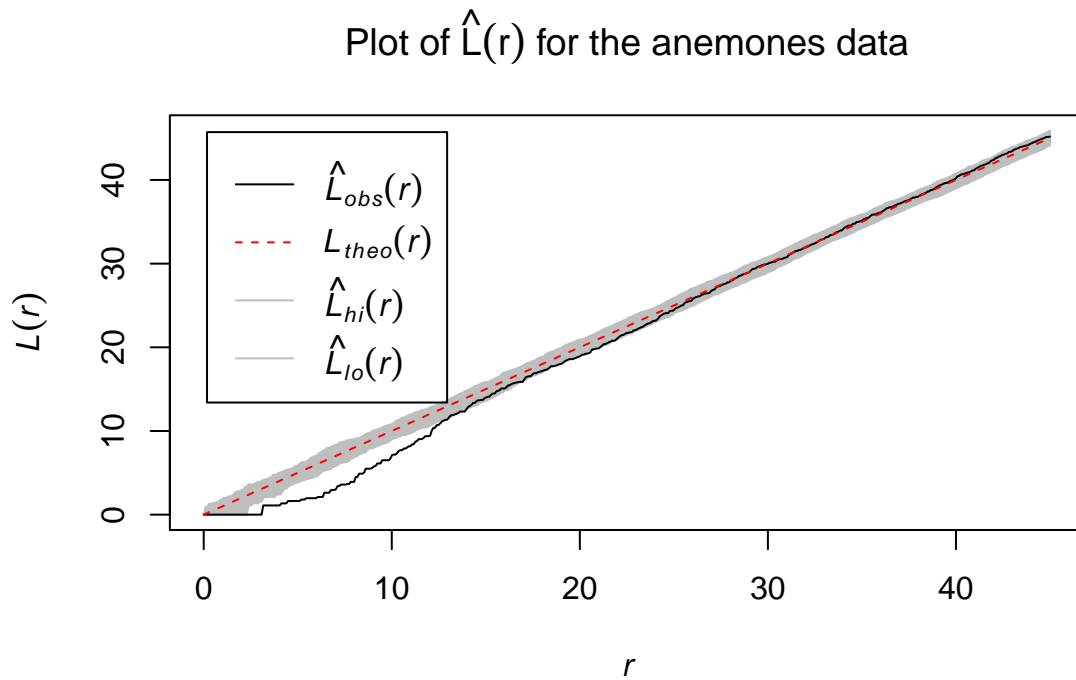


This averaged estimate is naturally much less variable than the estimate of any given realization. As the amount of simulation trials increases, the image plot would gradually become homogeneous in color, due to the estimate being more accurate per pixel. Also, the estimates that are farthest from the true value of λ are those near the edges; this could be ameliorated somewhat by specifying the `edges=` command inside `density.ppp`.

2. Plot data and use the $L(h)$ function to test CSR.

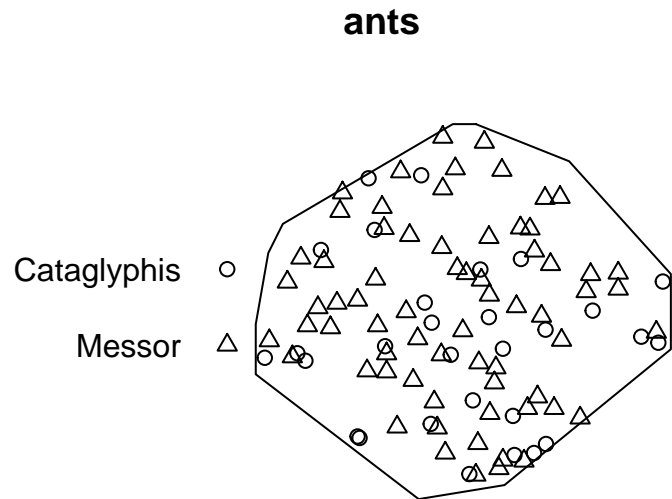


- First dataset, **anemones**:

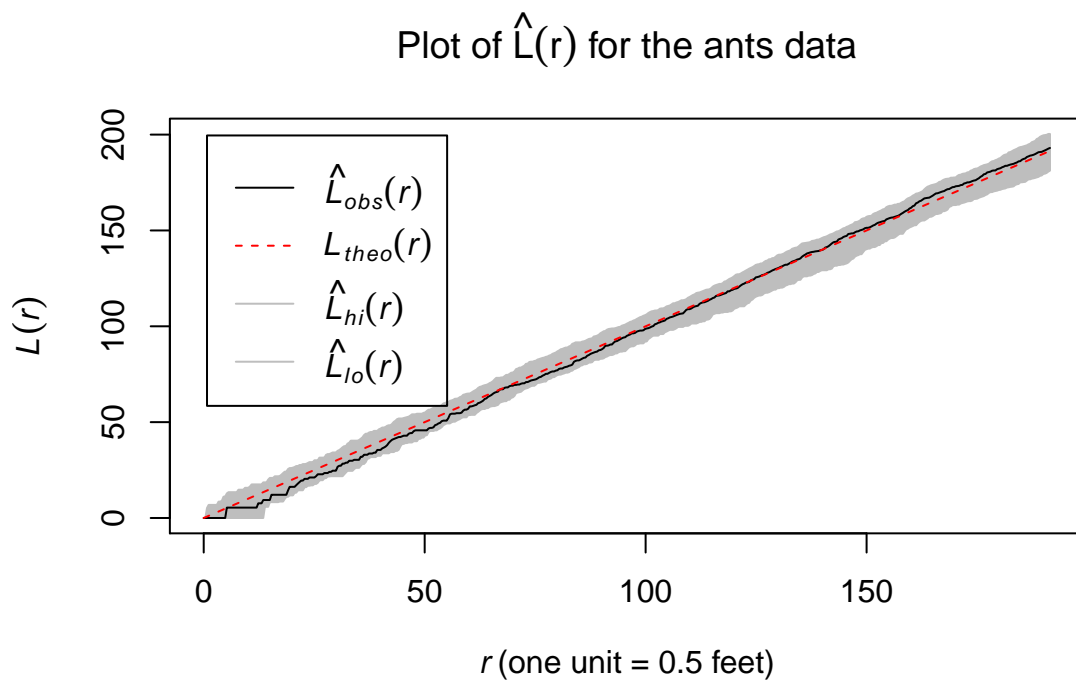


Based on the simulated datasets, the observed **anemones** data exhibits significantly fewer events at small values of h than would be expected under Complete Spatial Randomness. Once h is approximately greater than

30, the data follow CSR fairly closely, but for almost half of the data the observed $\hat{L}(h)$ is outside of the simulated envelopes, meaning there is evidence against CSR in these data.

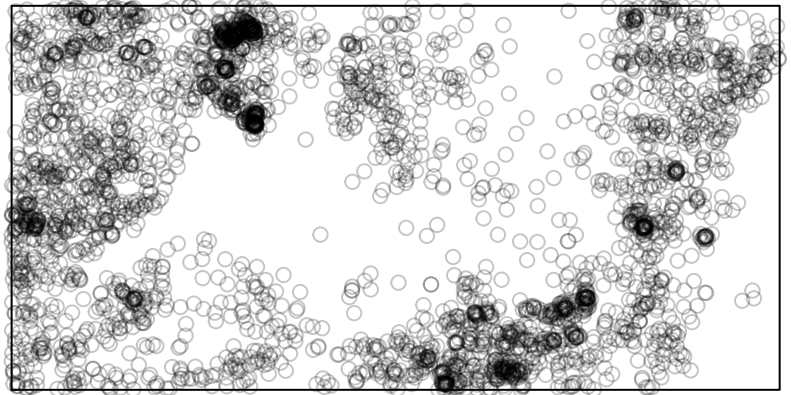


- Second dataset, **ants**:



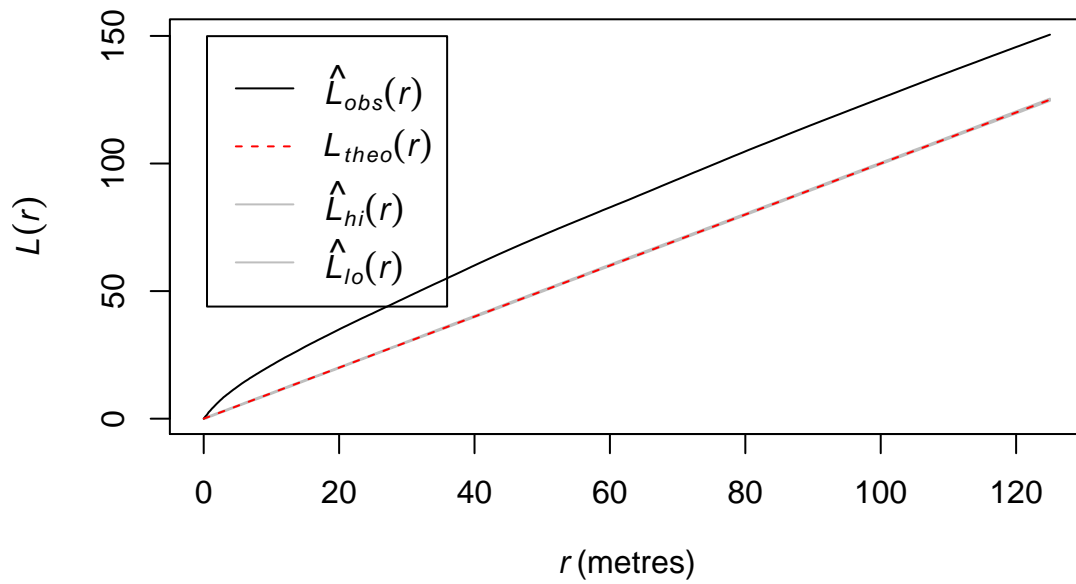
Based on the simulated datasets, there is no significant evidence that the **ants** were not generated by a homogeneous Poisson process.

bei



- Third dataset, **bei**:

Plot of $\hat{L}(r)$ for the rainforest data



As can be seen from the plot of the data, the events are clustered. Indeed, the L function indicates that at all distances there is a higher-than-expected number of events given CSR. Therefore, there is strong evidence that the rainforest data were not generated by a homogeneous Poisson process.