

Homework 4 - Stat 534
Due Friday, February 10, 2017

- For $\lambda = 30$ generate 9 realizations of *CSR* on the unit square. For each realization, construct a kernel estimate of $\lambda(s)$. How do the estimated intensity functions compare to the constant intensity under *CSR*? What precautions does this exercise suggest with regard to interpreting estimates of intensity from a single realization (or data set)? The following R code will simultaneously produce the data and plots.

```
par(mfrow=c(3,3))
for(i in 1:9) plot(density(rpoispp(30)))
```

Provide me with the images.

- In class we looked at a heterogeneous Poisson process on the unit square with intensity function

$$\lambda(x, y) = \exp(5x + 2y)$$

- Simulate a realization of the process using the following R code.

```
sim.dat<-rpoispp(function(x,y){exp(5*x + 2*y)})
```

Plot the results and comment.

- Plot simulation envelopes for the K function (or some suitable modification of it) and comment.
- Fit a trend model to your data using `ppm`. Provide me with the parameter estimates and associated standard errors.
- Check the fit using `quadrat.test`. Use `method="MonteCarlo"` instead of the large sample chi-squared test. Plot the results Discuss.
- Compare these results from this model with those to a model fit under an assumption of CSR. Summarize the results. Provide me the model comparison results (AIC comparisons are fine).
- Plot a nonparametric estimate of the intensity function. Compare the fitted surface you got using `ppm` with the nonparametric (kernel density estimate) surface in some suitable way.

- Recall the use of the `nncorr` statistic in the Finland Pines data set. The distribution of heights (the marks) was of interest. We saw that the nearest neighbor correlation between heights was -0.1839798 . We questioned whether or not this was unusual. Carry out a randomization test to assess this. You can use the `rlabel` command to scramble the marks if you want. Provide me with a histogram of the randomization distribution and a p-value. Discuss **BRIEFLY** your results. Provide me with your R-code, also. (Note - you need to extract the correlation from the `nncorr` output. Here is how to do that:

```
nncorr(finpines.ht)[ "correlation"]
correlation
-0.1839798
```

4. Let's derive a K function for something other than a CSR process. We will assume a Neyman-Scott process with the following properties.

- i The parent process is a homogeneous Poisson process with intensity λ .
- ii The number of offspring produced by each parent (N) is homogeneous Poisson with intensity μ .
- iii The position of each offspring is determined by a bivariate normal distribution with mean $(0, 0)$ (i.e. it is centered over the parent) and variance-covariance matrix $\sigma^2\mathbf{I}$. Note that this implies that the x and y coordinates are determined independently of one another with the same variance.

Consider 2 offspring from the same parent located at (X_1, Y_1) and (X_2, Y_2) .

- (a) What is the distribution of

$$W = \frac{(X_1 - X_2)^2}{2\sigma^2} + \frac{(Y_1 - Y_2)^2}{2\sigma^2}?$$

- (b) Note that the Euclidean distance between the 2 points is

$$H = [(X_1 - X_2)^2 + (Y_1 - Y_2)^2]^{1/2} = (2\sigma^2 W)^{1/2}.$$

Derive the cdf of H .

- (c) Recall that we were told that the K function for Neyman-Scott processes with homogeneous Poisson parent processes and radially symmetric $f(h)$ is

$$K(h) = \pi h^2 + \frac{E(N(N-1))}{\lambda E(N)^2} F(h).$$

Use the results from above to find the K function for the described process.