Lab 3

STAT 517 - Winter 2023

1. Lake Constance, bordered by Austria, Germany and Switzerland, is Europe's third largest lake. Records of freezes of major parts of the lake has been kept since the year 875, and a total of 38 major freezes have been reported since then and until 1976. The years in which freezes have occurred are provided in the file `constance_freezes.R`.

   (a) Fit a homogeneous Poisson process model to this data. Provide an estimate of the intensity of the process, as well as a 95% confidence interval for it.

   (b) Note both the first and last observations of the process are censored. How much do your results change if ignore this fact?

   (c) How can you use a quantile-quantile plot and/or a Kolmogorov???Smirnov test to determine whether the homogenous Poisson process is a good model for this data?

2. The dataset in the object `myscallops.R` shows the locations off the coast of New Jersey and Long Island where catches of scallops have exceeded 10 units during during a particular year. You can plot the data using the following R code

```
library(maps)
setwd("~/Documents/Courses/UW/STAT517/Lab 3")
load("./myscallops.R")
par(mar=c(1,1,1,1)+0.2)
map("usa", xlim=c(-74,-71), ylim=c(38.2,41.5), fill=TRUE)
points(myscallops10$long, myscallops10$lat, cex=0.75, pch=20)
```

   (a) Fit a non-homogenous Poisson process model to this data using a nonparametric model for the intensity function based on a Gaussian kernel density estimator.

   (b) Plot the contour lines of the intensity function on the map and discuss any potential shortcoming of your estimator you can see.

3. Consider a Hawkes process with intensity function:

$$\lambda(t \mid \mathcal{H}_t) = \mu + \sum_{t_i < t} \frac{k}{\{c + (t - t_i)\}^p}$$

Figure 1: True density and observations used to illustrate the EM algorithm for fitting a location and scale mixture of three bivariate Gaussian distributions.

(a) Assume that $\mu = 1$ $p = 2$, $c = 1$ and $k = 0.5$. Is this a stationary Hawkes process?

(b) Simulate of sample of this Hawkes process over the interval $[0, T]$ and plot the (realized) intensity function associated with it.

(c) Write the likelihood function for this sample and compute the maximum likelihood estimator of the various model parameters.