Math 365 Poisson Processes

Submit your R script to tleise@amherst.edu by next Tuesday.

The objective of this lab is to explore basic properties and behavior of Poisson processes.

1 Poisson process simulations

1. Run several simulations using the code posted on the course schedule page (PoissonProcessDemo.R) using different values of the rate parameter λ. Plot both a histogram of the interarrival times and a graph of the cumulative sum giving the arrival times for each simulation.

2. Visually compare the density function $f(t) = \lambda e^{-\lambda t}$ (plotted as a solid line in the demo script) to the histogram for the simulated interarrival times.

3. For each simulation, also compare the average number of arrivals per unit time during the simulation to the expected value λ (using the cumulative sum plot). Report several examples.

4. The step function plot showing number of customers who have arrived by time t should look roughly like a line with what slope?

2 Real life example

1. Download the file CoalMineAccidents.csv from the course webpage and then load into RStudio. You may want to set the folder where you save this file as your working directory (navigate to it in the Files tab, then under More select Set as Working Directory, or use the `setwd` command). The following code loads the data into RStudio as a data frame:

   ```
   f <- read.table(file="CoalMineAccidents.csv", sep=",", header=TRUE)
   ```

   This file has data on days between mining accidents in the UK. The first listed accident occurred on March 15, 1851, and the last one occurred on January 12, 1918. We can roughly model this data as a Poisson process.

2. Plot a histogram of the days between mining accidents to see if it looks roughly exponential:

   ```
   hist(f[,"Days"], freq=FALSE) # density plot
   ```

3. What would you estimate for λ? A good estimate can be made by simply calculating the average number of accidents per day from the data. Add the density function for your λ estimate to the histogram to see how well it seems to fit the data.

4. Plot the cumulative sum of the “Days” data. Does this look like a time-homogeneous process?

5. To improve the rate parameter estimate, break the data into two time periods, splitting at the point where the rate parameter appears to have significantly changed. Estimate λ for each time period. To specify a range like the first 100 accidents, type `f[1:100,"Days"]`.