Assignment #1 — Due Wednesday, January 28, 2008, by 5:00 P.M.

Turn in homework in lecture, discussion, or your TA's mailbox. Indicate the discussion section in which you expect to attend to pick up this assignment on the assignment.

311: Monday 1:20–2:10 **312:** Monday 12:05–12:55

The first several problems relate to the Poisson process as described in lecture and in section 11.6 of the textbook. The last problems are related to the molecular evolution example from lecture.

1. Let W_1, W_2, ... be independent and identically distributed (i.i.d) Exponential(2) random variables (with density $f(x) = 2e^{-2x}$ for $x \geq 0$). Define $S_0 = 0$ and let $S_n = \sum_{i=1}^{n} W_i$ be the sum of the first n W_i. Define $N_t = \max\{n : S_n \leq t\}$ for all $t \geq 0$. If the values $S_1, S_2, \ldots$ are plotted on a number line and considered to be *events*, then N_t counts the number of events in the interval $[0, t]$.

   (a) Let $Y = N_1$ be the number of events up to time one. Write the event $\{Y = 0\}$ as equivalent to an event involving W_1. Find the probability of this event both by making a calculation using the density of W_1 and by using the Poisson probability distribution.

   (b) Write the event $\{Y = 1\}$ as equivalent to an event involving W_1 and W_2. Find the probability of this event both by making a calculation using the joint density of W_1 and W_2 and by using the Poisson probability distribution.

   (c) Follow the R example in the file `poisson-process.R` to run a simulation of a Poisson process 100,000 times and find the empirical distribution for the number of events before time 1. Use this simulation as a numerical approximation of the probability function for Y.

   (d) Use the function `dpois()` in R to compute the exact Poisson probabilities $\mathsf{P}(Y = k)$ for $k = 0, 1, \ldots, 8$. How do these exact probabilities agree with the numerical estimates from the simulation of the Poisson process?

2. Define a Poisson process as in the previous problem with rate $\lambda = 2$. Define random variables $X_1 = N_2$, $X_2 = N_3 - N_1$, and $X_3 = N_4 - N_2$.

   (a) Which pairs of X_i are independent?

   (b) What is $\mathsf{E}(X_i)$ for $i = 1, 2, 3$?

   (c) What is $\mathsf{P}(X_1 = 2)$?

   (d) What is $\mathsf{P}(X_1 = 2 \mid X_2 = 1)$?

   (e) What is $\mathsf{P}(X_1 = 2 \mid X_3 = 1)$?

3. Consider the following probability model which is related to the molecular evolution example from lecture. You are interested in estimating a parameter t which is related to a parameter θ by the function $\theta = 1 - e^{-t}$. You have a statistical model that $X \sim \text{Binomial}(10, \theta)$ and observe $X = 3$. Find a reasonable estimate of t.

4. Using the same setting as the previous model, suppose that $X_i \sim \text{Binomial}(10, \theta)$ for $i = 1, \ldots, 5$ and that the $\{X_i\}$ are mutually independent. You observe these values: 3, 5, 1, 1, 3. Find a reasonable estimate of t.

Work to do, but not turn in.

- If you do not yet have R installed on your computer, follow the instructions to do so.
- Read Chapter 5 and Chapter 11, section 6 of the textbook.