Math 480a: Sage, Spring 2011, Homework 8

William Stein

Due May 25, 2011

Instructions: Do the following 2 problems, and turn them in by the beginning of class on Wednesday, May 25, 2011. If you get help from a classmate or friend on any homework problem, please explicitly thank them in your write up. (This is the last homework assignment of the course, and should be pretty straightforward: get to work on your final projects!)

1. Compute the median, mean and standard deviation of the following list of 100 integers:

$$1, 2, 3, 4, \ldots, 99, 100$$

   using each of the approaches listed below. Note that some will give exact answers and some floating point, and that's fine. Also, some will default to a biased estimate and some won't, so the standard deviation can be slightly different between them—just use whatever the default is in each case.

   (a) The builtin Sage commands `std`, `mean`, and `median`, which are immediately available at the Sage prompt when you startup Sage (no import required).

   (b) Using `rpy2` (see lecture 23 from May 18).

   (c) Using R via the pexpect interface, e.g., `r.eval` or `%r` in the notebook (see lecture 22 from May 16).

   (d) The `scipy.stats` module.

   (e) The `stats.TimeSeries` object.

   (f) Three functions that you write from scratch. [If you want, you may find copying from a certain lecture on Cython in which we implemented standard deviation in various ways helpful.]

2. Repeat Problem 1, but with the million numbers `[1..10^6]`, and time each. Which approach is the fastest and which is slowest? (Don't count the time to make the object in each case, but just the time to compute the relevant statistic.)