CS 485 Assignment 11, due April 28, 2006

1. If x and y are independent variables with expected value zero, prove that $\sigma^2(xy) = \sigma^2(x)\sigma^2(y)$.

2. Consider random sequence of length 1000 of integers 0 through 9. Represent a sequence by the set of its subsequences of length three. Represent a set of subsequences by a sketch consisting of the ten lowest subsequences in some random ordering of subsequences. Experimentally what is the resemblance of two sketches in terms of the resemblance of the original sequences? You might try sequences with 10 errors, 20 errors, etc.

3. Generate a stream of 10,000 characters from an alphabet of size 100. Count the number of occurrences of each element. Using the data stream technique of hashing each element to + or -1, estimate the number of occurrences of each element. Repeat the process seven times using different randomly selected hash functions.

4. Do a literature search for sketches and find an interesting algorithm or problem.