CS485 Spring 2007
Homework 13

Due Date: April 27 2007

NOTE: To speed up homework grading, please submit each homework problem on a separate sheet of paper, with you name and NetID on the top. Thank you!

1. Experiment with the "add 1 with probability $\frac{1}{2^k}$" method of counting number of occurrences of 0 in a binary stream. Give the mean and variance of the counter in a stream of length $10,000$ where the fraction of zeros is $0.1, 0.5$ and 0.9.

2. How much memory was used by the algorithm for answering the query "stream of length N has fewer than $\frac{T}{2}$ or more than $2T$ distinct elements" presented in class? How much memory is needed in order to assure correctness of the answer with 99% probability (that is: if A is the number of distinct elements in the stream, then if $A < \frac{T}{2}$ the algorithm should output "NO" with probability at least 99%, and if $A > \frac{2}{T}$, the algorithm should output "YES" with probability at least 99%)? How much memory is needed to answer it exactly (in terms of N and T)?

3. Prove the lemma used at the end of Lecture 35 in analyzing an alternative way for estimating the number of distinct elements in a datastream. The method finds the minimum element min of $H(x)$, where $H : \{1, \ldots, m\} \to \{1, \ldots, M = m^2\}$ is a hash function and x comes from the original datastream. The estimate is then $\hat{d} \approx \frac{M}{min}$. Finally, the lemma that you are asked to prove is: $\frac{d}{6} \leq \frac{M}{min} \leq 6d$ with probability at least $\frac{2}{3}$.

4. Let A be a symmetric real matrix, and let B be a symmetric invertible real matrix (i.e. there exists matrix B^{-1} such that $B^{-1}B = I$). Show that $\lambda_1 = \max_{||x||=1} \frac{x^T Ax}{x^T Bx}$ is the largest eigenvalue of $B^{-1}A$ (that is $(B^{-1}A)x = \lambda_1 x$ for some vector x).