15-451/651 Algorithms, Spring 2019
Recitation #12 Worksheet

Chebyshev's inequality

Chebyshev's inequality states that for a random variable X with expectation $\mathbf{E}[X]$ and variance $\mathbf{Var}[X]$, for any $\lambda > 0$, $\mathbf{Pr}[|X - \mathbf{E}[X]| \geq \lambda] \leq \frac{\mathbf{Var}[X]}{\lambda^2}$. Prove Chebyshev's inequality. Hint: use Markov's inequality.

Variance of the Sum. Suppose we take k independent copies of a random variable X with expectation $\mathbf{E}[X]$ and variance $\mathbf{Var}[X]$. Let the k copies be denoted $X_1, \ldots, X_k$. Let $T = \frac{1}{k}\sum_{i=1}^{k} X_i$. Then $\mathbf{E}[T] = \mathbf{E}[X]$ by linearity of expectation. Argue that $\mathbf{Var}[T] = \frac{1}{k}\mathbf{Var}[X]$.

Flipping Coins. Suppose you flip n independent coins, each with heads probability p. Let X be the number of heads. What is the expectation $\mu = \mathbf{E}(X)$ of X? What is the variance $sigma^2 = \mathbf{Var}(X)$? What is the probability that the number of heads differs from its expectation by more than λ? For the case where $p = 1/2$, what is the probability that the number of heads you see lies outside $n/2 \pm 10\sqrt{n}$?

CountSketch

CountSketch probability. For a random CountSketch matrix S with k rows and a fixed vector x, we showed in lecture and last recitation that $\mathbf{E}[\|Sx\|^2] = \|x\|^2$ and $\mathbf{Var}[\|Sx\|^2] = O(\|x\|^4/k)$. Assuming this, for what value of k do we have $\mathbf{Pr}[|\|Sx\|^2 - \|x\|^2| \geq \epsilon\|x\|^2] \leq 1/10$? Big-oh notation for k is fine. Please justify your answer.

CountSketch vs. CountMin. How is the CountSketch data structure different than the earlier CountMin hashing data structure we saw in class?

Sketch and Solve. Here is the sketch-and-solve paradigm for the approximate regression problem of outputting an $x' \in \mathbb{R}^d$ for which $\|Ax' - b\|_2^2 \leq (1+\epsilon) \min_x \|Ax - b\|_2^2$ with probability at least 9/10.

1. Draw S from a $k \times n$ random family of matrices for a value $k = O(d^2/\epsilon^2)$.
2. Compute $S \cdot A$ and $S \cdot b$.
3. Output the solution x' to $\min_{x'} \|(SA)x - (Sb)\|_2$.

What is the overall running time of this algorithm? Note you will need to account for the time for computing $S \cdot A$ and $S \cdot b$, as well as the time to solve the smaller regression problem $\min_x \|SAx - Sb\|_2^2$. You can assume the columns of SA are linearly independent. You can also assume each row of A has at least one non-zero entry (otherwise you can throw out the row without affecting the objective function). Assume A is represented in a way that the non-zero entries are stored in a list so can be accessed without reading the zero entries.