Sublinear Time and Space Algorithms 2026A – Problem Set 1

Robert Krauthgamer

Due: November 19, 2025

General instructions: Please keep your answers short and easy to read. You can use results, calculations or notation seen in class without repeating them, unless asked explicitly to redo them.

1. Prove that the storage requirement of Algorithm Morris+ seen in class is, with high constant probability, $O(\epsilon^{-2} \log \log m)$ bits.

   Hint: Use Markov's inequality.

2. Recall Algorithm FM seen in class (for estimating the number of distinct elements), and let Z denote the final value of z. Prove that $\mathbb{E}[Z^2] = \frac{2}{(d^*+1)(d^*+2)}$.

3. Design a streaming algorithm that at any time step m (not known in advance) receives a query $S \subset [n]$ and outputs an estimate of what fraction of items in the stream belong to S, within additive error $\epsilon \in (0, 1)$. Note that S is given only at query time (not in advance).

   Hint: Maintain $O(1/\epsilon^2)$ random samples and use them to estimate the fraction in S.