

CS5330: Hints for Week 3 Assignment

Assignment Due: Tuesday, 11th Feb 2020.

These are some hints for Assignment 3.

1. No hints :)
2. Let X_{ij} be the indicator random variable that balls i and j collide. What is $\mathbb{E}[X_{ij}]$ in terms of the probability vector? Let $X = \sum_{1 \leq i < j \leq m} X_{ij}$. Use Chebyshev's inequality to upper-bound $\Pr[X = 0]$ by $\text{Var}[X]/(\mathbb{E}[X])^2$. Compute $\mathbb{E}[X]$ and $\text{Var}[X]$. Show that if¹ $m \geq c/\|p\|_2$ for some constant c , then there will be a collision with probability $> 1/2$.
3. Let X_{ij} indicated that the i 'th and j 'th element (in the sorted order) are compared by the algorithm. We have already seen in the analysis of RandQS in the lecture that $X_{ij} = 1$ with probability $\frac{1}{j-i+1}$ and 0 otherwise. The number of comparisons is $X = \sum_{1 \leq i < j \leq n} X_{ij}$. You need to expand $(X - \mathbb{E}[X])^2$ and analyze $\mathbb{E}[X_{ij}X_{k\ell}] - \mathbb{E}[X_{ij}]\mathbb{E}[X_{k\ell}]$ in terms of how i, j, k, ℓ relate to each other.

¹ $\|p\|_2 = \sqrt{\sum_i p_i^2}$