Marburg University

Department of Computer Science
Prof. Dr. Sebastian Wild

Date: 2025-07-02
Version: 2025-07-02 14:59

Exercise Sheet 8 for Advanced Algorithms, Summer 2025

Hand In: *Until 2025-07-09 18:00, on ILIAS.*

Problem 1

40 points

We consider again the setting of throwing n balls into n bins using *pairwise-independent* bin choices. In this problem, you will construct a scenario where the fullest bin has occupancy $\Omega(\sqrt{n})$, showing that the $O(\sqrt{n})$ upper bound from class is tight up to constant factors if all we know about the bin choices is pairwise independence.

Specifically, let $B_1, \ldots, B_n \in [n]$ be the random (indices) chosen by the n balls. Show that there is a distribution for $(B_1, \ldots, B_n)$ with the following properties

1. The marginal distributions are all uniform, $B_i \overset{\mathcal{D}}{=} \mathcal{U}([n])$.
2. For any two balls $i \neq j$, the r.v. B_i and B_j are pairwise independent, i.e., $\forall p, q \in [n] : \mathbb{P}[B_i = p \wedge B_j = q] = \frac{1}{n^2}$
3. $\mathbb{E}[\max X_j] = \Omega(\sqrt{n})$

Hint: Try choosing a (fixed, but randomly selected) bin that each ball falls into with probability $1/\sqrt{n}$.

Problem 2

30 points

A sequence of random variables $Y_1, Y_2, \ldots \in U$ is called *k-wise independent* if for every k-tuple $(Y_{i_1}, \ldots, Y_{i_k})$ it holds that

$$\forall y_1, \ldots, y_k \in U \ : \ \mathbb{P}[(Y_{i_1}, \ldots, Y_{i_k}) = (y_1, \ldots, y_k)] \quad = \quad \mathbb{P}[Y_{i_1} = y_1] \cdot \cdots \cdot \mathbb{P}[Y_{i_k} = y_k]$$

A family of hash functions $\mathcal{H}$ is called k-wise independent if, for h uniformly chosen from $\mathcal{H}$, the random variables $h(x_1), \ldots, h(x_n)$ are k-wise independent.

Generalize the proof from class about the upper bound for the fullest-bin occupancy for the case the bins are picked by a hash function chosen randomly from a 3-wise independent family. What can you prove about the occupancy?