Stefan Walzer, Stefan Hermann
ITI Prof. Sanders
winter term 2025/2026

Exercise Sheet 7 – Classic Hash Tables

Probability and Computing

Exercise 1 – 2-Independence vs. 1-Universality

Let $\mathcal{H} \subseteq [m]^D$ be a family of hash functions mapping D to $[m]$. Prove or disprove the following implications:

(a) $\mathcal{H}$ is 2-independent $\Rightarrow$ $\mathcal{H}$ is 1-universal.

(b) $\mathcal{H}$ is 1-universal $\Rightarrow$ $\mathcal{H}$ is 2-independent.

Hint: In one case, the implication is straightforward. In the other, trivial counterexamples exist.

Exercise 2 – d-Independence without Mutual Independence

Alice and Bob each spin a roulette wheel with 10 equally sized segments labeled 0 to 9. Let A and B denote Alice's and Bob's outcomes, respectively. Define $C = (A + B) \bmod 10$.

(a) Show that A, B, and C are pairwise independent.

(b) Show that A, B, and C are not mutually independent.

(c) For any $d \in \mathbb{N}$, construct a family of random variables that is d-independent but not fully independent.

Exercise 3 – Find the Error

Let p be prime, $\mathbb{F}_p = \{0, \ldots, p-1\}$ and $m \in \mathbb{N}$. Consider the following class of hash functions from $\mathbb{F}_p$ to $[m]$, also mentioned in the lecture.

$$\mathcal{H} = \{x \mapsto ((a \cdot x) \bmod p) \bmod m \mid a \in \mathbb{F}_p^*\}.$$

Consider the following argument that $\mathcal{H}$ is 1-universal. Find the mistake in the proof.

The proof considers arbitrary $x, y \in \mathbb{F}_p$ with $x \neq y$. It has six steps.

$$\begin{aligned}
\Pr_{h\sim\mathcal{H}}[h(x) = h(y)] &\overset{1}{=} \Pr_{a\sim\mathcal{U}(\mathbb{F}_p^*)}[(ax \bmod p) \bmod m = (ay \bmod p) \bmod m] \\
&\overset{2}{=} \Pr_{a\sim\mathcal{U}(\mathbb{F}_p^*)}[((ax \bmod p) - (ay \bmod p)) \bmod m = 0] \\
&\overset{3}{=} \Pr_{a\sim\mathcal{U}(\mathbb{F}_p^*)}[((ax - ay) \bmod p) \bmod m = 0] \\
&\overset{4}{=} \Pr_{a\sim\mathcal{U}(\mathbb{F}_p^*)}[(a(x-y) \bmod p) \bmod m = 0] \\
&\overset{5}{=} \Pr_{u\sim\mathcal{U}(\mathbb{F}_p^*)}[u \bmod m = 0] \\
&\overset{6}{=} \frac{|\{m, 2m, 3m, \dots,\} \cap \mathbb{F}_p^*|}{|\mathbb{F}_p^*|} \\
&\overset{7}{\leq} \frac{1}{m}.
\end{aligned}$$

In Step 5 we use that the function $a \mapsto az \bmod p$ is a bijection on $\mathbb{F}_p^*$ for any fixed $z \in \mathbb{F}_p^*$. Therefore, if $a \sim \mathcal{U}(\mathbb{F}_p^*)$ and $u := az$ then $u \sim \mathcal{U}(\mathbb{F}_p^*)$.

Exercise 4 – Bonus: Concentration Bounds for Sums of d-wise Independent Random Variables

Let $d \in \mathbb{N}$ be even, and $\{X_1, \dots, X_n\}$ be a d-wise independent family of random variables, each distributed as $\mathrm{Ber}(p)$ with $p = \Omega(1/n)$.

Define $X = \sum_{i=1}^n X_i$. Note: X is not necessarily binomially distributed since the X_i are not mutually independent.

The goal is to prove the concentration bound: for any $\delta > 0$,

$$\Pr[X - \mathbb{E}[X] \geq \delta\mathbb{E}[X]] = O(\delta^{-d}(np)^{-d/2}).$$

To this end, consider the "centered" random variables $Y_i := X_i - p$, their sum $Y = \sum_{i=1}^n Y_i$, and the moment $\mathbb{E}[Y^d]$.

(i) Warm-up: Let $d \geq 3$ and $n \geq 3$. Verify and briefly explain why the following hold:

  (a) $\mathbb{E}[Y_1^5 Y_2^{42}] = \mathbb{E}[Y_1^5]\mathbb{E}[Y_2^{42}]$

  (b) $\mathbb{E}[Y_1^5 Y_2^{42} Y_3] = 0$

  (c) $\mathbb{E}[Y_1^5] \leq \mathbb{E}[Y_1^2]$

  In subsequent steps, you may apply these insights without further justification.

(ii) Show: $\mathbb{E}[Y_1^2] \leq p$.

(iii) Let $i_1, \dots, i_d \in [n]$ (not necessarily distinct) and $S = \{i_1, \dots, i_d\}$. Prove:

- If $|S| > d/2$, then $\mathbb{E}[Y_{i_1} \cdots Y_{i_d}] = 0$.
- Otherwise, $\mathbb{E}[Y_{i_1} \cdots Y_{i_d}] \leq p^{|S|}$.

(iv) Show: $\mathbb{E}[Y^d] = O((np)^{d/2})$. You may assume $d = O(1)$. **Hint:** Expand $(\sum_{i=1}^n Y_i)^d$. Yes, this yields n^d terms.

(v) Prove the original goal by applying Markov's inequality to Y^d.