1. **4-wise independence**

We saw in class a construction of 2-wise independent hash functions. Here, we will generalize this to 4-wise independence; a similar idea extends to k-wise independence for any k.

Let \( \mathbb{F} \) be a finite field. Define the following family of functions \( h : \mathbb{F} \to \mathbb{F} \):

\[
H = \{ h_{a,b,c,d} (x) = a + bx + cx^2 + dx^3 : a, b, c, d \in \mathbb{F} \}
\]

Prove that \( H \) is a 4-wise independent family. That is: for any 4 distinct values \( x_1, x_2, x_3, x_4 \in \mathbb{F} \), and any 4 values, not necessarily distinct, \( y_1, y_2, y_3, y_4 \in \mathbb{F} \),

\[
\Pr_{h \in H} [h(x_1) = y_1, h(x_2) = y_2, h(x_3) = y_3, h(x_4) = y_4] = \frac{1}{|\mathbb{F}|^4}
\]

2. **Sketching:**

Let \( U \) be a finite universe. A family of hash functions \( H = \{ h : U \to \{-1,1\} \} \) is called a sketching hash family with error \( \varepsilon \) if it satisfies the following property: for any labeling of the universe elements with real values \( v : U \to \mathbb{R} \), not all zero, it holds that

\[
\Pr_{h \in H} \left[ \sum_{x \in U} v(x)h(x) = 0 \right] \leq \varepsilon.
\]

Equivalently, we can identify each \( h \in H \) with the vector \( v_h \in \{-1,1\}^{|U|} \) given by the truth table of \( h \) (namely, its evaluation on all universe elements). The condition that \( H \) is a sketching hash family is equivalent to the property that any nonzero vector \( v \in \mathbb{R}^{|U|} \) is orthogonal to at most an \( \varepsilon \)-fraction of the vectors \( \{v_h : h \in H\} \).

The value \( \text{Sketch}_h(v) = \sum v(x)h(x) \) is called the sketch of \( v \), under the hash function \( h \).

(a) Prove that if \( v, v' \) are different labelings of the universe, then with high probabilities over the choice of the hash function \( h \), their sketches are distinct:

\[
\Pr_{h \in H} [\text{Sketch}_h(v) = \text{Sketch}_h(v')] \leq \varepsilon
\]

(b) Prove that if \( H \) is the set of all functions from \( U \) to \{-1,1\} then it is a sketching hash family with error \( \varepsilon = 1/2 \).

(c) Prove that if \( H \) is 4-wise independent then it is a sketching hash family with error \( \varepsilon = 2/3 \).

Hint: Fix a nonzero \( v : U \to \mathbb{R} \) and let \( \sigma^2 = \sum_{x \in U} v(x)^2 > 0 \). Define \( Q(h) = |\text{Sketch}_h(v)|^2 \). Prove that \( \mathbb{E}_{h \in H}[Q(h)] = \sigma^2 \) and \( \mathbb{E}_{h \in H}[Q(h)^2] \leq 3 \sigma^4 \). Use the Paley-Zygmund inequality (Google it) to prove that \( \Pr_{h \in H} [Q(h) > 0] \geq 1/3 \).
3. Satisfiability

We saw in class an algorithm for solving Satisfiability for 3-CNFs on \( n \) inputs, which runs in time \( 2^{\gamma n} \) for \( \gamma \approx 0.41 \), and improves upon full enumeration which takes time \( 2^n \).

The same ideas can be generalized to \( k \)-CNFs for \( k \geq 3 \).

(a) Show that the same algorithm, when applied to \( k \)-CNFs, finds a solution (if one exists) in expected time \( 2^{\gamma_m n} \) for some \( \gamma_m < 1 \).

(b) Show that \( \gamma_m \leq 1 - c/k \) for some absolute constant \( c > 0 \).

4. Random walks

Let \( X_0 = 0, X_1, X_2, ... \) be an random walk defined as \( X_i = X_{i-1} + \Delta_i \), where \( \Delta_i \in \{-1,1\} \) are independently chosen with probability \( \Pr[\Delta_i = 1] = \Pr[\Delta_i = -1] = 1/2 \). We will show that with a constant probability \( |X_n| \approx \sqrt{n} \), and that the same holds even if the steps \( \Delta_i \) are not independent, but are only 4-wise independent.

(a) We saw in class that \( \mathbb{E}[X_n^2] = n \). Show that by Markov’s inequality, this implies that

\[
\Pr[|X_n| \geq 4\sqrt{n}] \leq \frac{1}{16}
\]

(b) Show that

\[
\Pr\left[|X_n| \geq \frac{1}{2}\sqrt{n}\right] \geq \frac{n}{12}
\]

Hint: Compute \( \mathbb{E}[X_n^4] \) and use Paley-Zygmund.

(c) Conclude that \( \Pr\left[\frac{1}{2}\sqrt{n} \leq |X_n| \leq 4\sqrt{n}\right] \geq 1% \).

(d) Show that the same conclusion holds if the steps \( \Delta_i \) are only 4-wise independent, instead of being completely uniform.