Chapter 7

On k-wise independence

By Sariel Har-Peled, March 19, 2024[1]

7.1. Pairwise independence

7.1.1. Pairwise independence

Definition 7.1.1. A set of random variables $X_1, \ldots, X_n$ is ***pairwise independent***, if for any pair of values α, β, and any two indices i, j, we have that

$$\mathbb{P}\Big[X_i = \alpha \text{ and } Y_j = \beta\Big] = \mathbb{P}[X_i = \alpha]\,\mathbb{P}\Big[Y_j = \beta\Big].$$

Namely, the variables are independent if you look at pairs of variables. Compare this to the much stronger property of independence.

Definition 7.1.2. A set of random variables $X_1, \ldots, X_n$ is ***independent***, if for any t, and any t values $\alpha_1, \ldots, \alpha_t$, and any t indices $i_1, \ldots, i_t$, we have that

$$\mathbb{P}[X_{i_1} = \alpha_1, X_{i_2} = \alpha_2, \ldots, \text{ and } Y_{i_t} = \alpha_{i_t}] = \prod_{j=1}^{t} \mathbb{P}\Big[X_{i_j} = \alpha_j\Big].$$

7.1.2. A pairwise independent set of bits

Let n be a number which is a power of two. As such, $t = \log_2 n = \lg n$ is an integer. Let $X_0, \ldots, X_{t-1}$ be truly independent random bits, each one of them is 1 with probability $1/2$.

For a non-negative integer number x, let $\mathrm{bit}(x, j) \in \{0, 1\}$ be the jth bit in the binary representation of x. That is, we have $x = \sum_j \mathrm{bit}(x, j) 2^j$.

For an index $i = 1, \ldots, 2^t - 1$, we define $Y_i = \bigotimes_{j:\mathrm{bit}(i,j)=1} X_j$, where $\otimes$ is the *xor* operator.

Lemma 7.1.3. *The random variables $Y_1, Y_2, \ldots, Y_{n-1}$ are pairwise independent.*

Proof: We claim that, for any i, we have $\mathbb{P}[Y_i = 1] = \mathbb{P}[Y_i = 0] = 1/2$. So fix i, and let α be an index such that $\mathrm{bit}(i, \alpha) = 1$, and observe that this follows readily if pick all the true random variables $X_0, \ldots, X_{t-1}$ in such an order such that X_α is the last one to be set.

Next, consider two distinct indices i, i', and two arbitrary values v, v'. We need to prove that

$$\mathbb{P}[Y_i = v \text{ and } Y_{i'} = v'] = \mathbb{P}[Y_i = v]\,\mathbb{P}[Y_{i'} = v'] = \frac{1}{4}.$$

To this end, let $B = \{ j \mid \mathrm{bit}(i, j) = 1\}$ and $B' = \{ j \mid \mathrm{bit}(i', j) = 1\}$. If there is an index $\beta \in B \setminus B'$, then we have

$$\mathbb{P}[Y_i = v \mid Y_{i'} = v'] = \mathbb{P}\Bigl[\bigotimes_{j:\mathrm{bit}(i,j)=1} X_j = v \Bigm| Y_{i'} = v'\Bigr] = \mathbb{P}\Bigl[X_\beta \otimes \bigotimes_{j:\mathrm{bit}(i,j)=1} X_j = v \Bigm| Y_{i'} = v'\Bigr]$$
$$= \mathbb{P}\Bigl[X_\beta = \Bigl(v \otimes \bigotimes_{j:\mathrm{bit}(i,j)=1} X_j\Bigr) \Bigm| Y_{i'} = v'\Bigr] = \frac{1}{2}.$$

This implies that $\mathbb{P}[Y_i = v \text{ and } Y_{i'} = v'] = \mathbb{P}[Y_i = v \mid Y_{i'} = v']\, \mathbb{P}[Y_{i'} = v'] = (1/2)(1/2) = 1/4$, as claimed.

A similar argument implies that if there is an index $\beta \in B' \setminus B$, then $\mathbb{P}[Y_{i'} = v' \mid Y_i = v] = 1/2$, which implies the claim in this case.

Since $i \neq i'$, one of the two scenarios must happen, implying the claim. ■

7.1.3. An application: Max cut

Given a graph $\mathsf{G} = (\mathsf{V}, \mathsf{E})$ with n vertices and m edges, consider the problem of computing the max-cut. That is, computing the set of vertices S, such that the cut

$$(S, \overline{S}) = (S, \mathsf{V} \setminus S) = \{uv \in \mathsf{E} \mid u \in S, v \in \mathsf{V} \setminus S\}.$$

is of maximum cardinality.

7.1.3.0.1. Algorithm. To this end, let $Y_1, \ldots, Y_n$ be the pairwise independent bits of Section 7.1.2. Here, let S be the set of all vertices $v_i \in \mathsf{V}$, such that $Y_i = 1$. The algorithm outputs $(S, \overline{S})$ as the candidate cut.

7.1.4. Analysis

Lemma 7.1.4. *The expected size of the cut computed by the above algorithm is $m/2$, where $m = |\mathsf{E}(\mathsf{G})|$.*

Proof: Let Z_{uv} be an indicator variable for the event that the edge $uv \in \mathsf{E}$ is in the cut $(S, \overline{S})$.

We have that

$$\mathbb{E}\Bigl[\bigl|(S, \overline{S})\bigr|\Bigr] = \mathbb{E}\Bigl[\sum_{uv \in \mathsf{E}} Z_{uv}\Bigr] = \sum_{uv \in \mathsf{E}} \mathbb{E}[Z_{uv}] = \sum_{uv \in \mathsf{E}} \mathbb{P}[Y_u \neq Y_v] = |\mathsf{E}|/2,$$

using linearity of expectation and pairwise independence. ■

Lemma 7.1.5. *Given a graph G with n vertices and m edges, say stored in a read only memory, one can compute a max-cut of G, and the edges in it, using $O(\log n)$ random bits, and $O(\log n)$ RAM bits. Furthermore, the expected size of the cut is $\geq m/2$.*

Proof: The algorithm description is above. The pairwise independence is also described above, and requires only $O(\log n)$ random bits, which needs to be stored. Otherwise, all we need is to scan the edges of the graph, and for each one to decide if it is, or not in the cut. Clearly, this can be done using $O(\log n)$ RAM bits. ■

Compare this to the natural randomized algorithm of computing a random subset S. This would require using n random bits, and n bits of space to store it.

Max cut in the streaming model. Imagine that the edges of the graph are given to you via streaming: You are told the number of vertices in advance, but then edges arrive one by one. The above enables you to compute the cut in a streaming fashion using $O(\log n)$ bits. Alternatively, you can output the edges in a streaming fashion.

Another way of thinking about it, is that given a set $S = \{s_1, \ldots, s_n\}$ of n elements, we can use the above to select a random sample where every element is selected with probability half, and the samples are pairwise independent. The kicker is that to specify the sample, or decide if an element is in the sample, we can do it using $O(\log n)$ bits. This is a huge save compared to the regular n bits required as storage to remember the sample.

It is clear however that we want a stronger concept – where things are k-wise independent.

7.2. On k-wise independence

7.2.1. Definition

Definition 7.2.1. A set of variables $X_1, \ldots, X_n$ are ***k*-wise independent** if for any set $I = \{i_1, i_2, \ldots, i_t\}$ of indices, for $t \leq k$, and any set of values $v_1, \ldots, v_t$, we have that

$$\mathbb{P}[X_{i_1} = v_1 \text{ and } X_{i_2} = v_2 \text{ and } \cdots \text{ and } X_{i_t} = v_t] = \prod_{j=1}^{t} \mathbb{P}\big[X_{i_j} = v_j\big].$$

Observe, that verifying the above property needs to be done only for $t = k$.

7.2.2. On working modulo prime

Definition 7.2.2. For a number p, let $\mathbb{Z}_n = \{0, \ldots, n-1\}$.

For two integer numbers x and y, the ***quotient*** of x/y is $x \operatorname{div} y = \lfloor x/y \rfloor$. The ***remainder*** of x/y is $x \bmod y = x - y \lfloor x/y \rfloor$. If the $x \bmod y = 0$, than y ***divides*** x, denoted by $y \mid x$. We use $\alpha \equiv \beta \pmod{p}$ or $\alpha \equiv_p \beta$ to denote that α and β are ***congruent modulo p***; that is $\alpha \bmod p = \beta \bmod p$ – equivalently, $p \mid (\alpha - \beta)$.

Lemma 7.2.3. *Let p be a prime number.*

(A) For any $\alpha, \beta \in \{1, \ldots, p-1\}$, we have that $\alpha\beta \not\equiv 0 \pmod{p}$.

(B) For any $\alpha, \beta, i \in \{1, \ldots, p-1\}$, such that $\alpha \neq \beta$, we have that $\alpha i \not\equiv \beta i \pmod{p}$.

(C) For any $x \in \{1, \ldots, p-1\}$ there exists a unique y such that $xy \equiv 1 \pmod{p}$. The number y is the **inverse** *of x, and is denoted by x^{-1} or $1/x$.*

Proof: (A) If $\alpha\beta \equiv 0 \pmod{p}$, then p must divide $\alpha\beta$, as it divides 0. But α, β are smaller than p, and p is prime. This implies that either $p \mid \alpha$ or $p \mid \beta$, which is impossible.

(B) Assume that $\alpha > \beta$. Furthermore, for the sake of contradiction, assume that $\alpha i \equiv \beta i \pmod{p}$. But then, $(\alpha - \beta)i \equiv 0 \pmod{p}$, which is impossible, by (A).

(C) For any $\alpha \in \{1, \ldots, p-1\}$, consider the set $L_\alpha = \{\alpha * 1 \bmod p, \alpha * 2 \bmod p, \ldots, \alpha * (p-1) \bmod p\}$. By (A), zero is not in L_α, and by (B), L_α must contain $p-1$ distinct values. It follows that $L_\alpha = \{1, 2, \ldots, p-1\}$. As such, there exists exactly one number $y \in \{1, \ldots, p-1\}$, such that $\alpha y \equiv 1 \pmod{p}$. ■

Lemma 7.2.4. *Consider a prime p, and any numbers $x, y \in \mathbb{Z}_p$. If $x \neq y$ then, for any $a, b \in \mathbb{Z}_p$, such that $a \neq 0$, we have $ax + b \not\equiv ay + b \pmod{p}$.*

Proof: Assume $y > x$ (the other case is handled similarly). If $ax+b \equiv ay+b \pmod{p}$ then $a(x-y) \pmod{p} = 0$ and $a \neq 0$ and $(x-y) \neq 0$. However, a and $x-y$ cannot divide p since p is prime and $a < p$ and $0 < x-y < p$.■

Lemma 7.2.5. *Consider a prime p, and any numbers $x, y \in \mathbb{Z}_p$. If $x \neq y$ then, for each pair of numbers $r, s \in \mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, such that $r \neq s$, there is exactly one unique choice of numbers $a, b \in \mathbb{Z}_p$ such that $ax + b \pmod{p} = r$ and $ay + b \pmod{p} = s$.*

Proof: Solve the system of equations

$$ax + b \equiv r \pmod{p} \qquad \text{and} \qquad ay + b \equiv s \pmod{p}.$$

We get $a = \frac{r-s}{x-y} \pmod{p}$ and $b = r - ax \pmod{p}$. ■

7.2.3. Construction of k-wise independence variables

7.2.4. Construction

Consider the following matrix, aka the ***Vandermonde matrix***, defined by n variables:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \ldots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{bmatrix}.$$

Claim 7.2.6. $\det(V) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$.

Proof: One can prove this in several ways, and we include a proof via properties of polynomials. The determinant $\det(V)$ is a polynomial in the variables $x_1, x_2, \ldots, x_n$. Formally, let Π be the set of all permutations of $[\![n]\!] = \{1, \ldots, n\}$. For a permutation $\pi \in \Pi$, let $\operatorname{sign}(\pi) \in \{-1, +1\}$ denote the sign of this permutation. We have that

$$f(x_1, x_2, \ldots, x_n) = \det(V) = \sum_{\pi \in \Pi} \operatorname{sign}(\pi) x_i^{\pi(i)}.$$

Every monomial in this polynomial has total degree $\sum_{i=1}^n \pi(i) = 1 + 2 + \cdots + n = n(n-1)/2$. Observe, that if we replace x_j by x_i, then we have $f(x_1, \ldots, x_i, \ldots, x_{j-1}, x_i, x_{j+1}, \ldots, x_n)$ is the determinant of a matrix with two identical rows, and such a matrix has a zero determinate. Namely, the polynomial f is zero if $x_i = x_j$. This implies that $x_j - x_i$ divides f. We conclude that the polynomial $g \equiv \prod_{1 \leq i < j \leq n} (x_j - x_i)$ divides f. Namely, we can write $f = g * h$, where h is some polynomial.

Consider the monomial $x_2 x_3^2 \cdots x_n^{n-1}$. It appears in f with coefficient 1. Similarly, it generated in g by selecting the first term in each sub-polynomial, that is $\prod_{1 \leq i < j \leq n} \Big(x_j - x_i \Big)$. It is to verify that this is the only time this monomial appears in g. This implies that $h = 1$. We conclude that $f = g$, as claimed. ■

Claim 7.2.7. *If $x_1, \ldots, x_n$ are distinct, then the Vandermonde matrix V is invertible.*

Proof: By Claim 7.2.6, the determinant of V is $\det(V) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$. This quantity is non-zero if the xs are distinct, and a matrix is invertible in such a case. ■

Lemma 7.2.8. *For a vector* $\mathsf{b} = (b_0, \ldots, b_{k-1}) \in \mathbb{Z}_p^k$, *consider the associated polynomial* $f(x, \mathsf{b}) = \sum_{i=0}^{k-1} b_i x^i \bmod p$. *For any k distinct values $\alpha_1, \ldots, \alpha_k \in \mathbb{Z}_p$, and k values $v_1, \ldots, v_k \in \mathbb{Z}_p$, there is a unique choice of* b, *such that $f(\alpha_i) = v_i \bmod p$, for $i = 1, \ldots, k$.*

Proof: Let $\alpha_i = \left(1, \alpha_i, \alpha_i^2, \cdots, \alpha_i^{k-1}\right)$. We have that $f(\alpha_i, \mathsf{b}) = \langle \alpha_i, \mathsf{b} \rangle \bmod p$. This translates into the linear system

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix} \mathsf{b}^T = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} \iff \mathsf{M}\,\mathsf{b}^T = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} \qquad \text{where} \qquad \mathsf{M} = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \dots & \alpha_2^{n-1} \\ 1 & \alpha_3 & \alpha_3^2 & \dots & \alpha_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^{n-1} \end{bmatrix}.$$

The matrix M is the Vandermonde matrix, and by Claim 7.2.7 it is invertible. We thus get there exists a unique solution to this system of linear equations (modulo p). ■

The construction. So, let us pick independently and uniformly k values $b_0.b_1, \dots, b_{k-1} \in \mathbb{Z}_p$, let $\mathsf{b} = (b_0, b_1, \dots, b_{k-1})$. $g(x) = \sum_{i=0}^{k-1} b_i x^i \bmod p$, and consider the random variables

$$Y_i = g(i), \qquad \forall i \in \mathbb{Z}_p.$$

Lemma 7.2.9. *The variables $Y_0, \dots, Y_{p-1}$ are uniformly distributed and k-wise independent.*

Proof: The uniform distribution for each Y_i follows readily by picking b_0 last, and observing that each such choice corresponds to a different value of Y_i.

As for the k-independence, observe that for any set $I = \{i_1, i_2, \dots, i_k\}$ of indices, for $t \leq k$, and any set of values $v_1, \dots, v_k \in \mathbb{Z}_p$, we have that the event

$$Y_{i_1} = v_1 \text{ and } Y_{i_2} = v_2 \text{ and } \cdots \text{ and } Y_{i_k} = v_k$$

happens only for a unique choice of b, by Lemma 7.2.8. But there are p^k such choices. We conclude that the probability of the above event is $1/p^k = \prod_{j=1}^{k} \mathbb{P}\left[Y_{i_j} = v_j\right]$, as desired. ■

We summarize the result for later use.

Theorem 7.2.10. *let p be a prime number, and pick independently and uniformly k values $b_0.b_1, \dots, b_{k-1} \in \mathbb{Z}_p$, and let $g(x) = \sum_{i=0}^{k-1} b_i x^i \bmod p$. Then the random variables*

$$Y_0 = g(0), \dots, Y_{p-1} = g(p-1).$$

are uniformly distributed in $\mathbb{Z}_p$ and are k-wise independent.

7.2.5. Applications of k-wide independent variables

7.2.5.1. Product of expectations

Lemma 7.2.11. *If $X_1, \dots, X_k$ are k-wise independent, then $\mathbb{E}[X_1 \cdots X_k] = \mathbb{E}[X_1] \cdots \mathbb{E}[X_k]$.*

Proof: Immediate. ■

7.2.5.2. Application: Using less randomization for a randomized algorithm

We can consider a randomized algorithm, to be a deterministic algorithm **Alg**(x, r) that receives together with the input x, a random string r of bits, that it uses to read random bits from. Let us redefine **RP**:

Definition 7.2.12. The class **RP** (for Randomized Polynomial time) consists of all languages L that have a deterministic algorithm **Alg**(x, r) with worst case polynomial running time such that for any input $x \in \Sigma^*$,

- $x \in L \implies$ **Alg**$(x, r) = 1$ for half the possible values of r.
- $x \notin L \implies$ **Alg**$(x, r) = 0$ for all values of r.

Let assume that we now want to minimize the number of random bits we use in the execution of the algorithm (Why?). If we run the algorithm t times, we have confidence 2^{-t} in our result, while using $t \log n$ random bits (assuming our random algorithm needs only $\log n$ bits in each execution). Similarly, let us choose two random numbers from $\mathbb{Z}_n$, and run **Alg**(x, a) and **Alg**(x, b), gaining us only confidence $1/4$ in the correctness of our results, while requiring $2 \log n$ bits.

Can we do better? Let us define $r_i = ai + b \mod n$, where a, b are random values as above (note, that we assume that n is prime), for $i = 1, \ldots, t$. Thus $Y = \sum_{i=1}^{t}$ **Alg**(x, r_i) is a sum of random variables which are pairwise independent, as the r_i are pairwise independent. Assume, that $x \in L$, then we have $\mathbb{E}[Y] = t/2$, and $\sigma_Y^2 = \mathbb{V}[Y] = \sum_{i=1}^{t} \mathbb{V}[$**Alg**$(x, r_i)] \leq t/4$, and $\sigma_Y \leq \sqrt{t}/2$. The probability that all those executions failed, corresponds to the event that $Y = 0$, and

$$\mathbb{P}\Big[Y = 0\Big] \leq \mathbb{P}\Big[\big|Y - \mathbb{E}[Y]\big| \geq \frac{t}{2}\Big] = \mathbb{P}\Big[\big|Y - \mathbb{E}[Y]\big| \geq \frac{\sqrt{t}}{2} \cdot \sqrt{t}\Big] \leq \frac{1}{t},$$

by the Chebyshev inequality. Thus we were able to "extract" from our random bits, much more than one would naturally suspect is possible. We thus get the following result.

Lemma 7.2.13. *Given an algorithm* **Alg** *in* **RP** *that uses* $\lg n$ *random bits, one can run it t times, such that the runs results in a new algorithm that fails with probability at most* $1/t$*, and uses only* $2 \lg n$ *random bits.*

7.3. Higher moment inequalities

The following is the higher moment variant of Chebychev inequality.

Lemma 7.3.1. *For a random variable X, we have that* $\mathbb{P}\Big[|X - \mathbb{E}[X]| \geq t\mathbb{E}\Big[|X - \mathbb{E}[X]|^k\Big]^{1/k}\Big] \leq \dfrac{1}{t^k}$

Proof: Setting $Z = |X - \mathbb{E}[X]|^k$, and raising the inequality by a power of k, we have

$$\mathbb{P}\Big[|X - \mathbb{E}[X]| \geq t\mathbb{E}[|X - \mathbb{E}[X]|^k]^{1/k}\Big] = \mathbb{P}\Big[Z^{1/k} \geq t\,\mathbb{E}[Z]^{1/k}\Big] = \mathbb{P}\Big[Z \geq t^k\,\mathbb{E}[Z]\Big] \leq \frac{1}{t^k},$$

by Markov's inequality. ■

The problem is that computing (or even bounding) the kth moment $M_k(X) = \mathbb{E}\Big[|X - \mathbb{E}[X]\,|^k\Big]$ is usually not easy. Let us do it for one interesting example.

Lemma 7.3.2. *Consider k be an even integer and let* $X_1, \ldots, X_n$ *be n random independent variables such that* $\mathbb{P}[X_i = -1] = \mathbb{P}[X_i = +1] = 1/2$*. Let* $X = \sum_{i=1}^{n} X_i$*. Then, we have*

$$\mathbb{P}\Big[|X| \geq \frac{tk}{2}\sqrt{n}\Big] \leq \frac{1}{t^k}.$$

Proof: Observe that $\mathbb{E}[X] = n\,\mathbb{E}[X_1] = 0$. We are interested in computing

$$M_k(X) = \mathbb{E}\Big[X^k\Big] = \mathbb{E}\Big[\Big(\sum_i X_i\Big)^k\Big] = \mathbb{E}\Big[\sum_{i_1=1}^{n}\dots\sum_{i_k=1}^{n} X_{i_1}X_{i_2}\cdots X_{i_k}\Big] = \sum_{i_1=1}^{n}\dots\sum_{i_k=1}^{n}\mathbb{E}[X_{i_1}X_{i_2}\cdots X_{i_k}] \tag{7.1}$$

Consider a term in the above summation, where one of the indices (say i_1) has a unique value among $i_1, i_2, \dots, i_k$. By independence, we have

$$\mathbb{E}[X_{i_1}X_{i_2}\cdots X_{i_k}] = \mathbb{E}[X_{i_1}]\,\mathbb{E}[X_{i_2}\cdots X_{i_k}] = 0,$$

since $\mathbb{E}[X_{i_1}] = 0$. As such, in the above all terms that have a unique index disappear. A term that does not disappear is going to be of the form

$$\mathbb{E}\Big[X_{i_1}^{\alpha_1}X_{i_2}^{\alpha_2}\dots X_{i_\ell}^{\alpha_\ell}\Big] = \mathbb{E}\Big[X_{i_1}^{\alpha_1}\Big]\mathbb{E}\Big[X_{i_2}^{\alpha_2}\Big]\dots\mathbb{E}\Big[X_{i_\ell}^{\alpha_\ell}\Big]$$

where $\alpha_i \geq 2$, and $\sum_i \alpha_i = k$. Observe that

$$\mathbb{E}[X_1^t] = \begin{cases} 0 & t \text{ is odd} \\ 1 & t \text{ is even.} \end{cases}$$

As such, all the terms in the summation of Eq. (7.1) that have value that is not zero, have value one. These terms corresponds to tuples $T = (i_1, i_2, \dots, i_k)$, such that the set of values $I(T) = \{i_1, \dots, i_k\}$ has at most $k/2$ values, and furthermore, each such value appears an even number of times in T (here $k/2$ is an integer as k is even by assumption). We conclude that the total number of such tuples is at most

$$n^{k/2}(k/2)^k.$$

Note, that this is a naive bound – indeed, we choose the $k/2$ values that are in $I(T)$, and then we generate the tuple T, by choosing values for each coordinate separately. We thus conclude that

$$M_k(X) = \mathbb{E}\Big[X^k\Big] \leq n^{k/2}(k/2)^k.$$

Since k is even, we have $\mathbb{E}\Big[X^k\Big] = \mathbb{E}\Big[|X|^k\Big]$, and by Lemma 7.3.1, we have

$$\mathbb{P}\bigg[|X| \geq \frac{tk}{2}\sqrt{n}\bigg] = \mathbb{P}\bigg[|X| \geq t\Big(n^{k/2}(k/2)^k\Big)^{1/k}\bigg] \leq \mathbb{P}\bigg[|X| \geq t\mathbb{E}\Big[|X|^k\Big]^{1/k}\bigg] \leq 1/t^k. \qquad \blacksquare$$

Corollary 7.3.3. *Consider k be an even integer and let $X_1, \dots, X_n$ be n random independent variables such that $\mathbb{P}[X_i = -1] = \mathbb{P}[X_i = +1] = 1/2$. For $X = \sum_{i=1}^n X_i$, and any k, we have $\mathbb{P}\Big[|X| \geq k\sqrt{n}\Big] \leq 1/2^k$.*

Observe, that the above proof did not require all the variables to be purely independent – it was enough that they are k-wise independent. We readily get the following.

Definition 7.3.4. Given n random variables $X_1, \dots, X_n$ they are ***k-wise independent***, if for any k of them (i.e., $i_1 < i_2, \dots, i_k$), and any k values $x_1, \dots, x_k$, we have

$$\mathbb{P}\Big[\bigcap_{\ell=1}^{k}(X_{i_\ell} = v_\ell)\Big] = \prod_{\ell=1}^{k}\mathbb{P}[X_{i_\ell} = v_\ell].$$

Informally, variables are k-wise independent, if any k of them (on their own) looks totally random.

Lemma 7.3.5. *Let $k > 0$ be an even integer, and let $X_1, \dots, X_n$ be n random independent variables, that are k-wise independent, such that $\mathbb{P}[X_i = -1] = \mathbb{P}[X_i = +1] = 1/2$. Let $X = \sum_{i=1}^n X_i$. Then, we have*

$$\mathbb{P}\Big[|X| \geq \frac{tk}{2}\sqrt{n}\Big] \leq \frac{1}{t^k}.$$

References

[MR95] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge, UK: Cambridge University Press, 1995.