CPSC 536N: Randomized Algorithms 2014-15 Term 2

Lecture 13

Prof. Nick Harvey *University of British Columbia*

1 k-wise independence

A set of events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ are called ***k*-wise independent** if for any set $I \subseteq \{1, \ldots, n\}$ with $|I| \leq k$ we have

$$\Pr\left[\, \wedge_{i \in I} \mathcal{E}_i \,\right] \;=\; \prod_{i \in I} \Pr\left[\, \mathcal{E}_i \,\right].$$

The term **pairwise independence** is a synonym for 2-wise independence.

Similarly, a set of discrete random variables $X_1, \ldots, X_n$ are called **k-wise independent** if for any set $I \subseteq \{1, \ldots, n\}$ with $|I| \leq k$ and any values x_i we have

$$\Pr\left[\, \wedge_{i \in I} X_i = x_i \,\right] = \prod_{i \in I} \Pr\left[\, X_i = x_i \,\right].$$

References: Dubhashi-Panconesi Section 3.4.

Claim 1 *Suppose $X_1, \ldots, X_n$ are k-wise independent. Then*

$$\mathrm{E}\left[\, \textstyle\prod_{i \in I} X_i \,\right] \;=\; \prod_{i \in I} \mathrm{E}\left[\, X_i \,\right] \qquad \forall I \text{ with } |I| \leq k.$$

Proof: For notational simplicity, consider the case $I = \{1, \ldots, k\}$. Then

$$\begin{aligned}
\mathrm{E}\left[\, \textstyle\prod_{i=1}^{k} X_i \,\right] &= \sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} \Pr\left[\, \wedge_{i=1}^{k} X_i = x_i \,\right] \cdot \prod_{i=1}^{k} x_i \\
&= \sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} \prod_{i=1}^{k} \Pr\left[\, X_i = x_i \,\right] \cdot x_i \qquad (k\text{-wise independence}) \\
&= \Big(\sum_{x_1} \Pr\left[\, X_1 = x_1 \,\right] \cdot x_1 \Big) \cdots \Big(\sum_{x_k} \Pr\left[\, X_k = x_k \,\right] \cdot x_k \Big) \\
&= \prod_{i=1}^{k} \mathrm{E}\left[\, X_i \,\right].
\end{aligned}$$

□

Example. To get a feel for pairwise independence, consider the following three Bernoulli random variables that are pairwise independent but not mutually independent. There are 4 possible outcomes of these three random variables. Each of these outcomes has probability 1/4.

X_1	X_2	X_3
0	0	0
0	1	1
1	0	1
1	1	0

They are certainly not mutually independent because the event $X_1 = X_2 = X_3 = 1$ has probability 0, whereas $\prod_{i=1}^{3} \Pr[\, X_i = 1\,] = (1/2)^3$. But, by checking all cases, one can verify that they are pairwise independent.

1.1 Constructing Pairwise Independent RVs

Let $\mathbb{F}$ be a finite field and $q = |\mathbb{F}|$. We will construct RVs $\{\, Y_u \,:\, u \in \mathbb{F} \,\}$ such that each Y_u is uniform over $\mathbb{F}$ and the Y_u's are pairwise independent. To do so, we need to generate only *two* independent RVs X_1 and X_2 that are uniformly distributed over $\mathbb{F}$. We then define

$$Y_u \;=\; X_1 + u \cdot X_2 \qquad \forall u \in \mathbb{F}. \tag{1}$$

Claim 2 *The random variables $\{\, Y_u \,:\, u \in \mathbb{F} \,\}$ are uniform on $\mathbb{F}$ and pairwise independent.*

Proof: We wish to show that, for any distinct RVs Y_u and Y_v and any values $a, b \in \mathbb{F}$, we have

$$\Pr[\; Y_u = a \,\wedge\, Y_v = b \;] \;=\; \Pr[\, Y_u = a\,] \cdot \Pr[\, Y_v = b\,] \;=\; 1/q^2. \tag{2}$$

This clearly implies pairwise independence. It also implies that they are uniform because

$$\Pr[\, Y_u = a\,] \;=\; \sum_{b \in \mathbb{F}} \Pr[\, Y_u = a \,\wedge\, Y_v = b\,] \;=\; 1/q.$$

To prove (2), we rewrite the event

$$\{Y_u = a \,\wedge\, Y_v = b\} \quad \text{as} \quad \{X_1 + u \cdot X_2 = a\} \cap \{X_1 + v \cdot X_2 = b\}\,.$$

The right-hand side is a system of linear equations:

$$\begin{pmatrix} 1 & u \\ 1 & v \end{pmatrix} \cdot \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \;=\; \begin{pmatrix} a \\ b \end{pmatrix}.$$

There is a unique solution $x_1, x_2 \in \mathbb{F}$ for this equation because $\det\left(\begin{smallmatrix} 1 & u \\ 1 & v \end{smallmatrix}\right) = v - u \neq 0$. The probability that $X_1 = x_1$ and $X_2 = x_2$ is $1/q^2$. □

References: Mitzenmacher-Upfal Section 13.1.3, Vadhan Section 3.5.2.

1.2 The finite field $\mathbb{F}_{2^m}$

The vector space $\mathbb{F}_2^m$. Often we talk about the boolean cube $\{0,1\}^m$, the set of all m-dimensional vectors whose coordinates are zero or one. We can turn this into a binary vector space by defining an addition operaton on these vectors: two vectors are added using coordinate-wise XOR. This vector space is called $\mathbb{F}_2^m$.

Example: Consider the vectors $u = [0,1,1]$ and $v = [1,0,1]$ in $\mathbb{F}_2^3$. Then $u + v = [1,1,0]$.

More arithmetic. Vector spaces are nice, but they don't give us all the arithmetic that we might like to do. For example, how should we multiply two vectors? Given our definition of vector addition, it is not so easy to define multiplication so that our usual rules of arithmetic (distributive law, etc.) will hold. But, it can be done. The main idea is to view the vectors as polynomials, and multiply them via polynomial multiplication.

Vectors as polynomials. Basically, we introduce a variable x then view a vector $u \in \mathbb{F}_2^m$ as a degree-$(m-1)$ polynomial $p_u(x) = \sum_{a=0}^{m-1} u_a x^a$. To multiply u and v, we actually multiply $p_u(x) \cdot p_v(x)$. The trouble is that the resulting polynomial could now have degree $2m-2$, so we can no longer think of it as an m-dimensional vector. To map it back to an m-dimensional vector, we will use polynomial division.

An irreducible polynomial. For every m, there is a polynomial q_m with coefficients in $\mathbb{F}_2$ and of degree m that is *irreducible* (cannot be factored). I am not aware of an explicit formula for such a q_m, just like there is no explicit formula for integers that are prime (cannot be factored). Anyways, a q_m can be found by brute-force search, so let's assume q_m is known. If we divide any polynomial by q_m, the remainder is a polynomial of degree at most $m-1$. This is how we can map polynomials back to $\mathbb{F}_2^m$.

Multiplication. We now complete our definition of the multiplication operation. To multiply u and v, we multiply $p_u(x) \cdot p_v(x)$, then we take the remainder when dividing by q_m. The result is a polynomial of degree at most $m-1$, say $p_w(x)$, so it corresponds to a vector $w \in \mathbb{F}_2^m$. It turns out that, with this definition of multiplication, all our usual rules of arithmetic are satisfied: associative law, commutative law, distributive law, etc.

The resulting algebraic structure is called the finite field $\mathbb{F}_{2^m}$.

References: Wikipedia.

1.3 Pairwise independent hashing

In the construction (1) of pairwise independent random variables, notice that we can compute Y_u easily given X_1, X_2 and u. To make this concrete, define the function $h = h_{X_1,X_2} : \mathbb{F} \to \mathbb{F}$ by

$$h(u) = X_1 + uX_2.$$

It is common to think of h as a ***hash function***, because it is a random-like function.

The connection is to hash functions is more compelling if we consider the case $\mathbb{F} = \mathbb{F}_{2^m}$. Then h maps $\mathbb{F}_{2^m}$ to $\mathbb{F}_{2^m}$. We saw above that $\mathbb{F}_{2^m}$ is just an algebraic structure on top of $\{0,1\}^m$, the set of all binary strings of length m. So the function h is actually a random function mapping $\{0,1\}^m$ to $\{0,1\}^m$. The pair (X_1, X_2) is the ***seed*** of this hash function.

This discussion proves the following theorem.

Theorem 3 *For any $m \geq 1$, there is a hash function $h = h_s : \{0,1\}^m \to \{0,1\}^m$, where the seed s is a bit string of length $2m$, such that*

$$\Pr_s\left[\, h_s(u) = v \,\wedge\, h_s(u') = v' \,\right] \;=\; 2^{-2m} \qquad \forall u, u', v, v' \in \{0,1\}^m \;\; \textit{with } u \neq u'.$$

The following trivial generalization is obtained by deleting some coordinates in the domain or the range.

Corollary 4 *For any $m, \ell \geq 1$, there is a hash function $h = h_s : \{0,1\}^m \to \{0,1\}^\ell$, where the seed s is a bit string of length $2 \cdot \max m, \ell$, such that*

$$\Pr_s\left[\, h_s(u) = v \,\wedge\, h_s(u') = v' \,\right] \;=\; 2^{-2\ell} \qquad \forall u, u' \in \{0,1\}^m, v, v' \in \{0,1\}^\ell \;\; \textit{with } u \neq u'.$$

References: Vadhan Section 3.5.2.

1.4 Example: Max Cut with pairwise independent RVs

Once again let's consider the Max Cut problem. We are given a graph $G = (V, E)$ where $V = \{0, \ldots, n-1\}$. Our previous algorithm picked mutually independent random variables $Z_0, \ldots, Z_{n-1} \in \{0, 1\}$, then defined $U = \{\, i \,:\, Z_i = 1 \,\}$.

We will instead use a pairwise independent hash function. Let $m = \lceil \log_2 n \rceil$. Pick $s \in \{0,1\}^{2m}$ uniformly at random. We use the hash function $h_s : \{0,1\}^m \to \{0,1\}$ given by Corollary 4 with $\ell = 1$. Define $Z_i = h_s(i)$. Then

$$\begin{aligned}
\mathrm{E}\,[\,|\delta(U)|\,] &= \sum_{ij \in E} \Pr[\ (i{\in}U \wedge j{\notin}U) \ \vee\ (i{\notin}U \wedge j{\in}U)\ ] \\
&= \sum_{ij \in E} \Pr[\ i{\in}U \wedge j{\notin}U\ ] + \Pr[\ i{\notin}U \wedge j{\in}U\ ] \\
&= \sum_{ij \in E} \Pr[\,Z_i\,]\Pr\left[\,\overline{Z_j}\,\right] + \Pr\left[\,\overline{Z_i}\,\right]\Pr[\,Z_j\,] \qquad \text{(pairwise independence)} \\
&= \sum_{ij \in E} \Big((1/2)^2 + (1/2)^2 \Big) \\
&= |E|/2
\end{aligned}$$

So the original algorithm works just as well if we make pairwise independent decisions instead of mutually independent decisions for placing vertices in U. The following theorem shows one advantage of making pairwise independent decisions.

Theorem 5 *There is a deterministic, polynomial time algorithm to find a cut $\delta(U)$ with $|\delta(U)| \geq |E|/2$.*

PROOF: We have shown that picking $s \in \{0,1\}^{2m}$ uniformly at random gives

$$\mathrm{E}_s\,[\,|\delta(U)|\,] \ \geq\ |E|/2.$$

In particular, there exists some particular $s \in \{0,1\}^{2m}$ for which the resulting cut $\delta(U)$ has size at least $|E|/2$.

We can use exhaustive search to try all $s \in \{0,1\}^{2m}$ until we find one that works. The number of trials required is $2^{2m} \leq 2^{2(\log_2 n + 1)} = O(n^2)$. This gives a deterministic, polynomial time algorithm. □

References: Mitzenmacher-Upfal Section 13.1.2, Vadhan Section 3.5.1.

2 Construction of k-wise independent random variables

We now generalize the pairwise independent construction of Section 1.1 to give k-wise independent random variables.

Let $\mathbb{F}$ be a finite field and $q = |\mathbb{F}|$. We start with mutually independent RVs $X_0, \ldots, X_{k-1}$ that are uniformly distributed over $\mathbb{F}$. We then define

$$Y_u \ =\ \sum_{a=0}^{k-1} u^a X_a. \qquad \forall u \in \mathbb{F}. \tag{3}$$

Claim 6 *The random variables $\{\ Y_u \,:\, u \in \mathbb{F}\ \}$ are uniform on $\mathbb{F}$ and k-wise independent.*

PROOF:(Sketch) The argument is similar to Claim 2. In analyzing

$$\Pr\left[\, Y_{i_1} = t_1 \ \wedge \ \cdots \ \wedge \ Y_{i_k} = t_k \,\right]$$

we obtain the system of equations:

$$\begin{pmatrix} 1 & i_1 & \cdots & i_1^{k-1} \\ 1 & i_2 & \cdots & i_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & i_k & \cdots & i_k^{k-1} \end{pmatrix} \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} \quad = \quad \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_k \end{pmatrix}.$$

This has a unique solution because the matrix on the left-hand side is a Vandermonde matrix. □

References: Vadhan Section 3.5.5.

2.1 k-wise independent hash functions

Our construction of pairwise independent hash functions from Section 1.3 generalizes to give k-wise independent hash functions. As before, we focus on the special case $\mathbb{F} = \mathbb{F}_{2^m}$.

Given a seed $s = (X_0, \ldots, X_{k-1})$, we define $h_s : \mathbb{F}_{2^m} \to \mathbb{F}_{2^m}$ by

$$h(u) \ = \ X_0 + uX_1 + \cdots + u^{k-1} X_{k-1},$$

which equals the random variable Y_u in (3).

Theorem 7 *For any $m \geq 1$, there is a hash function $h = h_s : \{0,1\}^m \to \{0,1\}^m$, where the seed s is a bit string of length km, such that*

$$\Pr_s\left[\, h_s(u_1) = t_1 \ \wedge \ h_s(u_k) = t_k \,\right] \ = \ 2^{-km}$$

for all distinct $u_1, \ldots, u_k \in \{0,1\}^m$ and all $v_1, \ldots, v_k \in \{0,1\}^m$.

As before we obtain a trivial but useful generalization by shrinking the domain or range.

Corollary 8 *For any $m, \ell \geq 1$, there is a hash function $h = h_s : \{0,1\}^m \to \{0,1\}^\ell$, where the seed s is a uniformly random bit string of length $k \cdot \max\{m, \ell\}$, such that*

$$\Pr_s\left[\, h_s(u_1) = v_1 \ \wedge \ h_s(u_k) = v_k \,\right] \ = \ 2^{-k\ell}$$

for all distinct $u_1, \ldots, u_k \in \{0,1\}^m$ and all $v_1, \ldots, v_k$ in $\{0,1\}^\ell$.