Mathematical Toolkit **Autumn 2015**

Lecture 11: November 4, 2015

Lecturer: Madhur Tulsiani

1 Basics of probability: random variables

Recall that in the last lecture, we discussed how to define probability for infinite spaces using σ-algebras. In this case, we can define probability for any event which is a set in the σ-algebra. To define a random variable, we will also consider a σ-algebra $\mathcal{F}'$ on the range of the random variable, and allow only functions for which we can correctly define the probability of the random variable taking values in a given set in $\mathcal{F}'$.

Definition 1.1 *Let $\mathcal{F} \subseteq 2^{\Omega}$ and $\mathcal{F}' \subseteq 2^{\mathcal{R}}$ be σ-algebras. As function $f : \Omega \to \mathcal{R}$is said to be* measurable *(under the σ-algebras $\mathcal{F}$ and $\mathcal{F}'$) if for all $E' \in \mathcal{F}'$, we have $f^{-1}(E') \in \mathcal{F}$.*

We then define a require (say) a real-valued random variable to be a *measurable* function from Ω to $\mathbb{R}$. Unless otherwise specified, we consider the σ-algebra on the target space $\mathbb{R}$ to be the Borel σ-algebra. We will define the expectation of a random variable $X : \Omega \to \mathbb{R}$ as the integral with respect to the measure.

$$\mathbb{E}[X] \;=\; \int_{\Omega} X(\omega) d\mu \,.$$

The definition of the integral with respect to a measure with respect to a measure requires some amount of care, though we will not be able to discuss this in much detail.

2 Randomized polynomial identity testing

We use the above to prove the following lemma, which gives an algorithm for testing if a polynomial f in n variables $x_1, \ldots, x_n$ over a field $\mathbb{F}$ is identically zero.

Lemma 2.1 (Schwartz-Zippel lemma) *Let $f(x_1, x_2, \ldots, x_n)$ be a non-zero polynomial of degree $d \geq 0$, i.e.,*

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) \;&=\; \sum c_{i_1 i_2 \ldots i_n} \cdot x_1^{i_1} \cdot x_2^{i_2} \cdots x_n^{i_n} \\ s.t., \qquad & i_1 + i_2 + \ldots + i_n \leq d \end{aligned}$$

over a field, $\mathbb{F}$. Let $S \subseteq \mathbb{F}$, be a finite subset and let $x_1, x_2, \ldots, x_n$ be selected randomly from S independently. Then,

$$\mathbb{P}[f(x_1, x_2, \ldots, x_n) = 0] \;\leq\; \frac{d}{|S|} \,.$$

Proof: We will prove this lemma by induction on n. This lemma can be proved simply by using conditional probability.

Base Case: n = 1
A non zero polynomial, $f(x_1)$ can have at most d roots. Hence, $\mathbb{P}[f(x_1) = 0] \leq \frac{d}{|S|}$.

Induction Step
Assume that the lemma holds for any polynomial in $n-1$ variables. We need to prove that it holds true for $f(x_1, x_2, \ldots, x_n)$. We can write f as:

$$f(x_1, x_2, \ldots, x_n) = x_1^k \cdot g(x_2, \ldots, x_n) + h(x_1, x_2, \ldots, x_n)$$

where, k is largest degree of x_1. Thus we have $0 < k \leq d$ (If $k = 0$ then we are already done). We also have the $\deg(g(x_2, \ldots, x_n)) \leq d - k$.

Now let us define two events.

$$E \equiv \{f(x_1, x_2, \ldots, x_n) = 0\} \quad \text{and} \quad F \equiv \{g(x_2, \ldots, x_n) = 0\}$$

We can then write,

$$\mathbb{P}[E] = \mathbb{P}[F] \cdot \mathbb{P}[E|F] + \mathbb{P}[\neg F] \cdot \mathbb{P}[E|\neg F] .$$

We now analyze each of the terms. By the induction hypothesis, we have

$$\mathbb{P}[F] = \mathbb{P}[g(x_2, \ldots, x_n) = 0] = \frac{d-k}{|S|} .$$

Also, fixing the values of $x_2 = a_2, \ldots, x_n = a_n$ such that $g(a_2, \ldots, a_n) \neq 0$, $f(x_1, a_2, \ldots, a_n)$ is a degree-k polynomial in x_1. Thus, using the base case, we get that

$$\mathbb{P}[E|\neg F] \leq \frac{k}{|S|} .$$

Bounding the other two probabilities by 1, we get that

$$\mathbb{P}[E] \leq \frac{d-k}{|S|} \cdot 1 + 1 \cdot \frac{k}{|S|} = \frac{d}{|S|}$$

as desired. ■

2.1 An application: bipartite perfect matching

Consider the following example which applied the Schwartz-Zippel lemma for testing if a given bipartite graph has a perfect matching. Given a bipartite graph, $G = (U, V, E)$ with $|U| = |V| = n$, we say that the graph has a perfect matching, if there exists a set $E' \subseteq E$ of n edges, with exactly one edge in E' being incident on every vertex of G.

Let us define the Tutte matrix A as

$$A_{ij} = \begin{cases} x_{ij} & \text{if (i, j)} \in E \\ 0 & else \end{cases}$$

Note that A is not necessarily symmetric. The determinant of A can be written as,

$$\mathrm{Det}(A) = \sum_{\pi:[n]\to[n]} \mathrm{sign}(\pi) \prod_{i=1}^{n} A_{i,\pi(i)}$$

where π defines the permutation from rows to columns. Note that the determinant is a degree-n polynomial in the variables x_{ij}. Verify the follwing:

Exercise 2.2 *G has a perfect matching if and only if Det$(A) \not\equiv 0$.*

In this case, computing the determinant is expensive with $n!$ terms. But if we are given the values of the variables x_{ij}, we can simply compute the determinant using the Gaussian elimination method. The degree of the polynomial above is n. Thus, if we assign all variables randomly from a set of $2n$ real values, if $\mathrm{Det}(A) \not\equiv 0$, we will detect it with probability at least $1/2$.

The randomized algorithm by Schwartz-Zippel Lemma can be used to parallelize the checking as well. There is no known deterministic algorithm for this problem which can be parallelized efficiently.

3 Computing expectations

We will demonstrate the computation of expectations with some examples.

Let Z be a random variable of number of heads associated with n tosses of coins. Let X_i be the random variable associated with the i^{th} toss, defined as

$$X_i = \begin{cases} 1 & \text{if toss i is heads} \\ 0 & \text{if toss i is tails} \end{cases} .$$

Thus $\Omega = \{0,1\}^n$. Let us assume that the coin tosses are independent of each other (though, as you will see, we will not need this assumption here).

Example 3.1 *With the assumption that $\mathbb{P}[X_i = 1] = \mathbb{P}[X_i = 0] = 1/2$, we want to compute $\mathbb{E}[Z]$.*

Since $Z = \sum X_i$, we have, $\mathbb{E}[Z] = \sum \mathbb{E}[X_i]$. Now,

$$\begin{aligned} \mathbb{E}[X_i] &= 1 \cdot \mathbb{P}[X_i = 1] + 0 \cdot \mathbb{P}[X_i = 0] \\ &= \mathbb{P}[X_i = 1] = 1/2 \end{aligned}$$

Hence $\mathbb{E}[Z] = n/2$.

Note that if a random variable, X_e takes a binary value if an event, e occurs or not, then the expected value of X_e is $\mathbb{P}[e]$.

Example 3.2 *Instead of the uniform probability of heads and tails, if $\mathbb{P}[X_i = 1] = p$ and $\mathbb{P}[X_i = 0] = 1 - p$, then $\mathbb{E}[Z] = n \cdot p$.*

Note that we did not use independence in the above calculations. We just needed that for each i, $\mathbb{E}X_i = p$.

For the next example, we consider an *infinite* sequence of independent coin tosses, with $\mathbb{P}[\text{heads}] = p$ for each coin.

Example 3.3 *Given, that $\mathbb{P}[heads] = p$, what is $\mathbb{E}[\#tosses\ till\ the\ first\ heads]$?*

We define Z as the number of tosses till the first heads. Let E be the event that the first toss is heads. Then we have,

$$\begin{aligned} \mathbb{E}[Z] &= \mathbb{E}[Z|E] \cdot \mathbb{P}[E] + \mathbb{E}[Z|\neg E] \cdot \mathbb{P}[\neg E] \\ &= 1 \cdot \mathbb{P}[E] + (1 + \mathbb{E}[Z]) \cdot (1 - p) \end{aligned}$$

Thus we have, $\mathbb{E}[Z] = \dfrac{1}{p}$.

The above is known as a *geometric random variable* with parameter p.