COMP8601: Advanced Topics in Theoretical Computer Science
Lecture 1: Probabilistic Method, Markov's Inequality, Chebyshev's Inequality
Instructor: Hubert Chan
Date: 3 Sept 2013

These lecture notes are supplementary materials for the lectures. They are by no means substitutes for attending lectures or replacement for your own notes!

1 Probabilistic Method

1.1 Possible Events

The first principle of probabilistic method captures the trivial observation that if an event happens with positive probability, then it is possible for it to happen.

Fact 1.1 *Suppose $A \subset \Omega$ is an event such that $Pr(A) > 0$. Then, $A \neq \emptyset$. In particular, there exists $\omega \in \Omega$ such that $\omega \in A$.*

This principle is often used to prove existence of structures having certain properties.

1.1.1 Mono-chromatic Sets

Consider a finite set U and subsets $S_1, S_2, \ldots, S_m$ of U such that each S_i has size $|S_i| = l$. Is it possible to color each element of U *red* or *blue* such that no set S_i contains elements with only one color?

Proposition 1.2 *Suppose $m < 2^{l-1}$. Then, it is possible to color each element of U red or blue such that for all i, the set S_i contains elements from 2 colors.*

Proof: We run the following experiment. For each element, we color it red with probability $\frac{1}{2}$ and blue with probability $\frac{1}{2}$. This is performed independently over all points. The sample space Ω is the set of all possible colorings.

For each i, let A_i be the event that the set S_i contains elements of only one color. We wish to show that $Pr(\cap_{i=1}^m \overline{A_i}) > 0$. Hence, it suffices to show that $Pr(\cup_{i=1}^m A_i) = 1 - Pr(\cap_{i=1}^m \overline{A_i}) < 1$.

Observe that the event A_i happens means that all elements in S_i are all red or all blue. Hence, $Pr(A_i) = \frac{1}{2^l} + \frac{1}{2^l} = \frac{1}{2^{l-1}}$.

By the union bound, $Pr(\cup_{i=1}^m A_i) \leq \sum_{i=1}^m Pr(A_i) = m \cdot \frac{1}{2^{l-1}} < 1$. ■

1.2 Random Variables and Expectation

The next principle states that if it is possible for a random variable to take values at least as large as its mean.

Fact 1.3 *Suppose $E[X] = x$. Then, $Pr[X \geq x] > 0$.*

This principle is used for showing that there exists solution having objective value at least some certain number.

1.2.1 Max Cut

Suppose $G = (V, E)$ is a graph. A cut $C \subset V$ is a subset of V. An edge $e = \{u, v\} \in E$ is in cut C if $e \cap C = 1$. The edges in a cut is $E(C) := \{e \in E : e \cap C = 1\}$. The problem of Max Cut is to find a cut C such that the number $|E(C)|$ of cut edges is maximized.

Here is a very simple randomized algorithm. We form a random subset C in the following manner. Independently for each vertex v in V, we assign it a number 0 or 1, each with probability $\frac{1}{2}$. Then, the cut C consists of the vertices with number 1.

Proposition 1.4 $E[|E(C)|] = \frac{|E|}{2}$

Proof:

For each edge $e \in E$, let Y_e be the random variable that takes value 1 if $e \in E(C)$ and 0 otherwise. Then, $|E(C)| = \sum_{e \in E} Y_e$.

Consider an edge $e = \{u, v\}$. Note that $Y_e = 1$ *iff* exactly 1 of $\{u, v\}$ is in C, i.e. either (1) $u \in C$ and $v \notin C$ or (2) $u \notin C$ and $v \in C$.

Hence, it follows that $Pr(Y_e = 1) = \frac{1}{2}$. Therefore, $E[Y_e] = \frac{1}{2}$. By the linearity of expectation, $E[|E(C)|] = \sum_{e \in E} E[Y_e] = \frac{|E|}{2}$. ∎

1.2.2 Max 3-SAT

Suppose $x_0, x_1, \ldots, x_{n-1}$ are n Boolean variables. Consider a 3-CNF formula (Conjunctive Normal Form) with m clauses: $C_1 \wedge C_2 \wedge \cdots \wedge C_m$, where each C_j is a disjunction of 3 literals from 3 different variables. A literal is either a variable (e.g. x_1) or its negation (e.g. $\neg x_1$). A clause is satisfied if at least one of its 3 literals evaluates to TRUE.

Given a 3-CNF formula, the goal is to find an assignment of the variables so that as many clauses as possible are satisfied. Here is a randomized procedure for finding an assignment. Independently for each variable x_i, assign its value to be TRUE or FALSE, each with probability $\frac{1}{2}$.

Proposition 1.5 *The expected number of satisfied clauses is* $\frac{7m}{8}$.

Proof: Let Y_j be a random variable that takes value 1 if the clause C_j is satisfied and 0 otherwise. Then, the number of satisfied clauses is $\sum_{j=1}^{m} Y_j$. Observe that exactly 3 variables are included in the clause C_j, and out of the $2^3 = 8$ possible configurations for those 3 variables, exactly 1 configuration causes all 3 literals to be FALSE. Hence, $E[Y_j] = \frac{7}{8}$, and so $E[\sum_{j=1}^{m} Y_j] = \sum_{j=1}^{m} E[Y_j] = \frac{7m}{8}$. ∎

2 Markov's Inequality and Chebyshev's Inequality

Theorem 2.1 (Markov's Inequality) *Suppose X is a random variable taking non-negative values.*

For all $\alpha > 0$, $Pr(X \geq \alpha) \leq \frac{E[X]}{\alpha}$.

Remark. Since $X > \alpha$ implies that $X \geq \alpha$, we also have $Pr[X > \alpha] \leq \frac{E[X]}{\alpha}$.

Proof:

$E[X] = Pr(X \geq \alpha) \cdot E[X|X \geq \alpha] + Pr(X < \alpha) \cdot E[X|X < \alpha]$.

Observe that $E[X|X \geq \alpha] \geq \alpha$, $Pr(X < \alpha) \geq 0$ and $E[X|X < \alpha] \geq 0$.

Hence, $E[X] \geq Pr(X \geq \alpha) \cdot \alpha$. Rearranging gives the result. ∎

Theorem 2.2 (Chebyshev's Inequality) *Suppose X is a random variable with expectation $E[X] = \mu$. Then, for all $\alpha > 0$, $Pr[|X - \mu| \geq \alpha] \leq \frac{var[X]}{\alpha^2}$, where $var[X] = E[(X - \mu)^2]$.*

Proof: Observe that $Pr[|X - \mu| \geq \alpha] = Pr[(X - \mu)^2 \geq \alpha^2]$. Hence, by Markov's Inequality, $Pr[(X - \mu)^2 \geq \alpha^2] \leq \frac{E[(X-\mu)^2]}{\alpha^2} = \frac{var[X]}{\alpha^2}$. ∎

2.1 Comparing Markov and Chebyshev

We show that we can often obtain a better result using Chebyshev's inequality, if we have a good bound on the variance of the random variable involved.

Consider the example of flipping a fair coin 1000 times. We want to find an upper bound for the event that there are at least 600 heads. Let X be the number of heads. Then, $\mu = E[X] = 500$ and $var[X] = 250$.

Using Markov's Inequality, we have $Pr[X \geq 600] \leq \frac{500}{600} = \frac{5}{6}$.

Observe that $X \geq 600$ implies that $|X - \mu| \geq 100$. Hence, using Chebyshev's Inequality, we have $Pr[X \geq 600] \leq Pr[|X - \mu| \geq 100] \leq \frac{250}{100^2} = 0.025$.