Additive Combinatorics

Hui Sun

February 1, 2024

Contents

Chapter 1

Probabislitic Methods

1.1 First Moment Method

The first moment method upperbounds the probabilty of events occuring using expected value.

Warning 1.1. The first moment method does not give a lower bound.

If we want to show A contains a subset B that satisfies property $\mathcal{P}$, then it suffices to show that a randomly chosen subset B satisfies $\mathcal{P}$ with positive probability.

Definition 1.1 (expected value, variance)**.** Let X be a real-valued random variable with discrete support, then its expected value is

$$\mathbb{E}(X) = \sum_x xP(X)$$

And its variance is given by

$$Var(X) = \mathbb{E}(|X - \mathbb{E}(X)|^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

If we let $I(E)$ denote the indicator function on E, $I(E) = \begin{cases} 1, E \text{ happens} \\ 0, E \text{ doesn't happen} \end{cases}$ Then $\mathbb{E}(I(E)) = P(E)$. Let F be an event with nonzero probability.

Definition 1.2 (conditional probability and expected value)**.** The conditional probability of event E with respect to F is

$$P(E|F) = \frac{P(E \wedge F)}{P(F)}$$

We define conditional expected value as follows

$$\mathbb{E}(X|F) = \frac{\mathbb{E}(XI(F))}{\mathbb{E}(I(F))} = \sum_x xP(X = x|F)$$

A random variable is called Boolean if it only takes values $\{0, 1\}$, equivalently, it is the indicator function of some event E.

Proposition 1.1 (Borel-Cantellilemma). Let $E_1, \ldots, E_n$ be a sequence of events (possibly infinite), such that $\sum_n P(E_n) < \infty$, then

$$P(\text{ fewer than } M \text{ events happen}) \geq 1 - \frac{\sum_n P(E_n)}{M}$$

Proof. By moving terms, we show the probability of more than M events happen is $\leq \frac{\sum_n P(E_n)}{M}$. By Markov's inequality,

$$P(\sum_n I(E_n) \geq M) \leq \frac{\mathbb{E}(\sum_n I(E_n))}{M} = \frac{\sum_n P(E_n)}{M}$$

Note that in the last equality, we need the assumption there are finitely many events, which we can do by monotone convergence $\sum_n P(E_n) < \infty$. □

1.2 Second Moment Method

The second moment method tells us that X cannot deviate from the expected value $\mathbb{E}(X)$ too much, and these tools are known as large deviation inequalities.

Theorem 1.2 (Chebyshev's inequality). Let X be a random variable, then for any $\lambda > 0$, we have

$$P(|X - \mathbb{E}(X)| > \lambda Var(X)^{1/2}) \leq \frac{1}{\lambda^2}$$

Proof. If $Var(X) = 0$, then $X = \mathbb{E}(X)$, hence the inequality satisfies. Let $Var(X) > 0$, we note

$$P(|X - \mathbb{E}(X)| > \lambda Var(X)^{1/2}) = P(|X - \mathbb{E}(X)|^2 > \lambda^2 Var(X))$$

And by Markov's inequality, we have

$$P(|X - \mathbb{E}(X)|^2 > \lambda^2 Var(X)) \leq \frac{\mathbb{E}(|X - \mathbb{E}(X)|^2)}{\lambda^2 Var(X)} = \frac{1}{\lambda^2}$$

□

From this, we know that $X = E(X) + O(Var(X)^{1/2})$ with high probability. And $Var(X) = \mathbb{E}(|X - \mathbb{E}(X)|^2)$, then we know $|X - \mathbb{E}(X)|^2 \geq Var(X)$ with postiive probability. The tools using variance is the second moment method.

Warning 1.3. The second moment method gives two-sided bounds on how X distributes, i.e., the two tails decay at $1/\lambda^2$.

Example 1.1. One can derive the genrealized Chebyshev's inequality as follows, using a generalized notion of variance.

$$P(|X - \mathbb{E}(X)|^p > \lambda^p \mathbb{E}(|X - \mathbb{E}(X)|^p)) \leq \frac{1}{\lambda^p}$$

Recall the linearity of expected value is as follows:

$$\mathbb{E}\left(\sum_n X_n\right) = \sum_n \mathbb{E}(X_n)$$

But this is not true for variance in general! If $\{X_j\}$ are pairwise independent random variables, then we have

$$Var\left(\sum_n X_n\right) = \sum_n Var(X_n)$$

Lemma 1.1. Let X_j be random variables, then we have

$$Var\left(\sum_n X_n\right) = \sum_n Var(X_n) + \sum_{i,j,i\neq j} Cov(X_i, X_j)$$

where $Cov(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$.

Let's do an example.

Example 1.2. Let B be a randomly chosen set in A, then we have

$$Var(|B|) = \sum_a P(a \in B) - P(a \in B)^2$$

Proof. We have $|B| = \sum_a 1_B(a)$, and $a_1, a_2 \in B$ are pairwise independent events. Then we have

$$Var(|B|) = \sum_a Var(1_B(a))$$

And $Var(1_B(a))) = \mathbb{E}(1_B(a)^2) - \mathbb{E}(1_B(a))^2$, hence $Var(|B|) = \sum_a P(a \in B) - P(a \in B)^2$. $\square$

We note that $Var(|B|) \leq \mathbb{E}(|B|)$.

$\neq$