# CS-541 Homework 2

Avirat Belekar

October 16, 2020

## 1 Introduction

We intuitively feel it is rare for an observation to deviate from its expected value. Inequalities like Markov and Chebyshev's place this intuition on a firm mathematical ground. Markov's inequality relate probabilities to expectation and provide bounds for the cumulative distribution function of a random variable. On the other hand, Chebyshev's inequality relates variance to probabilities for a distribution function of a random variable.

## 2 Markov's Inequality

The most elementary tail bound is Markov's inequality , which asserts that for a positive random variable X with finite mean,

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t} \tag{1}$$

Intuitively if the mean of a positive random variable is small then it is more likely to be too large often which implies that the probability that it is large is small. While Markov on its own is fairly crude it will form the basis for much more refined tail bounds.

### 2.1 Proof

Suppose X is a discrete random variable,

$$\mathbb{E}[X] = \sum_{x} x.Pr(X = x) \geq \sum_{x \geq t} x.Pr(X = x) \geq t \sum_{x \geq t} Pr(X = x) = t.Pr(X \geq t) \tag{2}$$

Rearranging the terms we get,

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t} \tag{3}$$

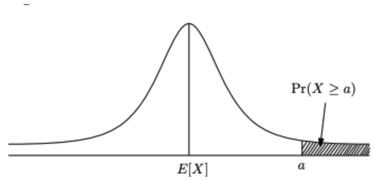


Figure 1: Markov's Inequality Bounds the probability of the shaded region

## 3 Chebyshev's Inequality

Chebyshev's inequality states that for a random variable X, with Var(X) $= \sigma^2$, for any $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]]| \geq t\sigma) \leq \frac{1}{t^2} \tag{4}$$

Before we prove this lets look at a simple application. We know that if average i.i.d random variables with mean $\mu$ and variance $\sigma^2$ , we have that the average :

$$\widehat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{5}$$

has mean $\mu$ and variance $\sigma^2/n$. So applying Chebyshev's inequality we obtain,

$$\mathbb{P}(|\widehat{\mu} - \mu| \geq \frac{t\sigma}{\sqrt{n}}) \leq \frac{1}{t^2} \tag{6}$$

So, with the probability at least 0.99 the average $10\sigma/\sqrt{n}$ from its expectation. This is pretty neat and almost directly gives us some thing called as Weak Law of Large Numbers.

### 3.1 Proof

Let us look at a picture that illustrates Chebyshev's Inequality. Higher moments often reveal more informa-

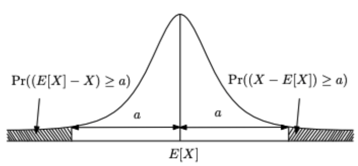


Figure 2: Chebyshev's Inequality Bounds the probability of the shaded region

tion about a random variable, which, in turn helps us derive better bounds. However, there is a trade-off. It is often difficult to compute higher moments in practical cases, e.g., while analyzing randomized algorithms. Now, let us look at the variance of a random variable.

$$Pr(|X - \mathbb{E}[X]| \geq a) = Pr((X - \mathbb{E}[X])^2 \geq a^2) = Pr(Y \geq a^2) \tag{7}$$

where $Y = (X - \mathbb{E}[X])^2$. Note that Y is a non-negative random variable. Therefore, using Markov's Inequality,

$$Pr(Y \geq a^2) \leq \frac{\mathbb{E}[Y]}{a^2} = \frac{\mathbb{E}(X - \mathbb{E}[X]^2)}{a^2} = \frac{Var[X]}{a^2} \tag{8}$$

## 4 Chernoff's Method

There are several refinements to the chebyshev's inequality. One simple one that is sometimes useful is to observe that if the random variable X has a finite k-th central moment then we can say that,

$$\mathbb{P}(|X - \mathbb{E}[X]]| \geq t) \geq \frac{\mathbb{E}|X - \mathbb{E}[X]|^k}{t^k} \tag{9}$$

For many random variables , the moment generating will exist in a neighbourhood around 0, i.e the moment generating function is finite for all $|t| \leq$ b where b $< 0$ is some constant. In these cases we can use the moment generating function to produce a tail bound. We define $\mu = \mathbb{E}[X]$. For an t $> 0$, we have that,

$$\mathbb{P}(|X - \mathbb{E}[X]]| \geq t) \geq \frac{\mathbb{E}|X - \mathbb{E}[X]|^k}{t^k} \tag{10}$$

by applying Markov's Inequality. Now that t is a parameter we can choose to get a tight upper bound, i.e. we can write this bound as :

$$\mathbb{P}((X-\mu) \geq u) = \inf_{0 \leq t \leq b} exp(-t(u+\mu))\mathbb{E}[exp(tX)] \tag{11}$$

This bound is known as Chernoff's Bound

# 5 Comparison between Chebyshev's Inequality and Markov's Inequality

Markov's inequality is a "large deviation bound". It states that the probability that a non-negative random variable gets values much larger than its expectation is small. Chebyshev's inequality is a "concentration bound". It states that a random variable with finite variance is concentrated around its expectation. The smaller the variance, the stronger the concentration.Both inequalities are used to claim that most of the time, random variables don't get "unexpected" values.A typical application of concentration bounds is polling - if you poll 200 people, then the standard deviation is so-and-so, and so the probability that the results are off by some amount is only which is very small. You can bound using Chebyshev's inequality, although there are better bounds. Another application is in the statistics of patients in a hospital. Suppose they conduct a survey and find that the average number of patients per day is 10. Suppose they want to be able all patients 90 percent of the time. Then they need to handle at most 100 patients. The list of probability results goes on. The law of large numbers states that the average of identical random variables tends (almost certainly) to the expectation.

# 6 Appendix

## 6.1 Weak Law of Large Numbers

Let $X_1, \ldots, X_n$ be independent random variables with the same mean $\mu$ and same variance $\sigma^2$ Define a new variable X such that,

$$X = \frac{X_1 + \ldots + X_n}{N} \tag{12}$$

then,

$$Pr[|X-\mu| \geq \alpha] \leq \frac{\sigma^2}{\alpha^2 N} \tag{13}$$

is that Weak of Law of Large Numbers