CS 535: Notes III

1 Three Bounds

Suppose we have two random variables X and Y which take the following values with equal probability.

$$X : 47, 48, 49, 50, 51, 52, 53$$

$$Y : 20, 30, 40, 50, 60, 70, 80$$

The expectation of both X and Y is 50. The values taken by Y are more spread out than the values taken by X. We capture this fact by *variance*. The *variance* of a random variable X is defined as follows.

$$Var(X) = E([X - E(X)]^2).$$

Thus variance is (square of) the average distance of X from its Expectation. In general, Variance is not linear. However, if the random variables are independent, then it is linear. If X and Y are two independent random variables, then $Var(X + Y) = Var(X) + Var(Y)$.

Given a random variable X, we are often interested in computing probabilities such as $\Pr(X > v)$, $\Pr(|X - E(X)| > a$. The following three inequalities help us to estimate this.

Markov's Inequality: Let X be a nonnegative random variable.

$$\Pr(X > v) \leq E(X)/v.$$

In other words,

$$\Pr(X > \alpha E(X)) \leq 1/\alpha.$$

Chebyshev's Inequality: Let X be random variable.

$$\Pr(|X - E(X)| \geq \delta) \leq Var(X)/\delta^2.$$

Chernoff's Bound: Let X_1, $X_2, \cdots, X_m$ be independent random variables that take values between 0 and 1. Let $E(X_1) = E(X_2) = \cdots = E(X_m) = p$. Let $X = X_1 + X_2 + \cdots + X_m$.

$$\Pr[|X/m - p| \geq p\delta] \leq 2e^{-\delta^2 mp/2}.$$

Suppose, we toss a fair coin n times. We would like to (upper) bound the probability that we see more than $2n/3$ heads. We do this using each of Markov, Chebyshev and Chernoff bounds. Let X denote the number of heads. We know that $E[X] = n/2$. Let us first calculate the variance of X. Let X_i be a random variable whose value if 1 if ith coin toss' outcome is head; otherwise

$X_i = 0$. Now $X = \sum i = 1^n X_i$. Since X_i's are independent, we have $Var(X) = \sum Var(X_i)$. Let us calculate $Var(X_i)$.

$$\begin{aligned}
Var(X_i) &= E([X_i - E(X_i)]^2) \\
&= E([X_i - \frac{1}{2}]^2) \\
&= E(X_i^2 - X_i + \frac{1}{4}) \\
&= E[X_i^2] - E[X_i] + \frac{1}{4} \\
&= \frac{1}{2} - \frac{1}{2} + \frac{1}{4} \\
&= \frac{1}{4}
\end{aligned}$$

Thus $Var(X) = n/4$.

Markov.

$$\begin{aligned}
\Pr[X > 2n/3] &\leq E[X] \times \frac{3}{2n} \\
&= \frac{n}{2} \times \frac{3}{2n} \\
&= \frac{3}{4}
\end{aligned}$$

Chebyshev.

$$\begin{aligned}
\Pr[X > 2n/3] &= \Pr[X - n/2 > n/6] \\
&\leq \Pr[|X - n/2| > n/6] \\
&= \Pr[|X - E[X]| > n/6] \\
&\leq \frac{Var(X)}{(n/6)^2} \\
&= \frac{9}{n}
\end{aligned}$$

Chernoff. We can write X as sum of $X_1, X_2, \cdots X_n$ which are independent, takes 0-1 values and have expectation 1/2. Thus, we can apply Chernoff bound.

$$\begin{aligned}
\Pr[X > 2n/3] &= \Pr[X - n/2 > n/6] \\
&\leq \Pr[|X - n/2| > n/6] \\
&= \Pr[|X/n - 1/2| > 1/6] \\
&= \Pr[|X/n - 1/2| > 1/2 \times 1/3] \\
&\leq 2e^{-(1/3)^2 \times n \times 1/2 \times 1/2} \\
&= 2e^{-n/36}
\end{aligned}$$