The Second-Moment Method

Will Perkins

January 28, 2013

Markov's Inequality

Recall Markov's inequality:

$$\Pr[|X| \geq t] \leq \frac{\mathbb{E}|X|}{t}$$

Proof:

$$\begin{aligned} \mathbb{E}|X| &= \int_{\omega:|X|<t} |X| \, dP + \int_{\omega:|X|\geq t} |X| \, dP \\ &\geq 0 + t \cdot \Pr[|X| \geq t] \end{aligned}$$

Markov's Inequality

What was special about the absolute value function?

1. non-negative
2. increasing

We can apply the same proof to other functions.

Chebyshev's Inequality

Theorem

Chebyshev's Inequality

$$\Pr[|X - \mathbb{E}X| \geq t] \leq \frac{var(X)}{t^2}$$

$$\mathrm{var}(X) = \mathbb{E}[|X - \mathbb{E}X|^2] \geq t^2 \cdot \Pr[|X - \mathbb{E}X| \geq t]$$

Chebyshev's Inequality

For example, if X has mean 3, variance 1, then the probability that X is more than 5 or less than 1 is bounded by $1/4$.

Is the bound a good bound for the Normal distribution?
Give an example of a random variable where the Chebyshev bound is tight.

The Weak Law of Large Numbers

As an application of Chebyshev's Inequality, we can prove our first limit theorem.

Theorem

Let $X_1, X_2, \ldots$ be i.i.d. random variables with mean μ and variance σ^2. Then for every $\epsilon > 0$,

$$\Pr\left[\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right] \to 0 \text{ as } n \to \infty$$

The Weak Law of Large Numbers

Comments:

- What does the WLLN say about political polling, for instance?
- Are all of the conditions necessary?
- Why do we say 'Weak'?

The Weak Law of Large Numbers

Proof: Let $U_n = \frac{X_1 + \ldots X_n}{n}$ and calculate $\mathbb{E}U_n$ and $\mathrm{var}(U_n)$

$$\mathbb{E}U_n = \mu$$

$$\mathrm{var}(U_n) = \frac{1}{n^2} \sum \mathrm{var}(X_i) = \frac{\sigma^2}{n}$$

Now apply Chebyshev:

$$\Pr\left[\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{\epsilon^2 n} \to 0 \text{ as } n \to \infty$$

First-Moment Method

Let X be a counting r.v. If $\mathbb{E}X \to 0$ as $n \to \infty$, then $\Pr[X = 0] \to 1$.

But what if $\mathbb{E}X \to \infty$? Can we say that $\Pr[X = 0] \to 0$?
No, not necessarily.
Example: Let $X = n^2$ with probability $1/n$ and 0 with probability $1 - 1/n$. Then $X = 0$ whp, but $\mathbb{E}X \to \infty$.

Counting Random Variables

If X is a counting random variable then by plugging in $t = \mathbb{E}X$,

Lemma

$$\Pr[X = 0] \leq \frac{var(X)}{(\mathbb{E}X)^2}$$

In particular, if $\mathbb{E}X \to \infty$ and $var(X) = o\left((\mathbb{E}X)^2\right)$, then $X \geq 1$ whp.

An Example

m balls thrown randomly into n bins. We saw with the first-moment method that if $m = (1+\epsilon)n \log n$, then whp there are no empty bins. But what if $m = (1-\epsilon)n \log n$?

Let X be the number of empty bins.

$$\mathbb{E}X = n \cdot \left(1 - \frac{1}{n}\right)^m$$

If $m = (1-\epsilon)n \log n$ then $\mathbb{E}X \sim n^\epsilon \to \infty$. But to conclude that $X \geq 1$ whp, we need the second-moment method.

Calculating the Variance

Let $X = X_1 + X_2 + \cdots + X_n$. Then

$$var(X) = \sum_{i=1}^{n} var(X_i) + \sum_{i \neq j} cov(X_i, X_j)$$

Let X_i be the indicator rv that the ith bin is empty.

Upper Bounding the Variance

Since we just need an upper bound to apply Chebyshev's inequality, things become simpler.

First, since the X_i's are indicator rv's, $var(X_i) \leq \mathbb{E}X_i$ and$\sum var(X_i) \leq \mathbb{E}X$. Since $\mathbb{E}X \to \infty$ for our choice of m,

$$\sum var(X_i) = o\left((\mathbb{E}X)^2\right)$$

Bounding the Covariances

We need to bound $\sum_{i \neq j} cov(X_i, X_j)$. Each of the terms are the same and there are $n(n-1)$ of them.

$$\begin{aligned} cov(X_i, X_j) &= \mathbb{E}(X_i X_j) - \mathbb{E}X_i \mathbb{E}X_j \\ &= \Pr[\text{ i and j empty}] - \Pr[\text{ i empty }] \cdot \Pr[\text{ j empty }] \\ &= \left(1 - \frac{2}{n}\right)^m - \left(1 - \frac{1}{n}\right)^{2m} \end{aligned}$$

We could calculate and show that this is small, but that is unnecessary: the covariance terms are negative.

Conclusion

We showed that for $m = (1 - \epsilon)n \log n$,

$$\mathbb{E}X \to \infty$$

and

$$var(X) = o\left((\mathbb{E}X)^2\right)$$

Then using Chebyshev's inequality with $t = \mathbb{E}X$, we concluded that $\Pr[X = 0] = o(1)$.

Applications of the 2nd Moment Method

- What's the 'typical' position of a simple random walk after n steps?
- What's the longest run of Heads in n flips of a fair coin?
- What is the maximum of n independent standard normal RV's?