Lecture 15: Markov and Chebyshev's Inequalities

Anup Rao

May 3, 2019

We talk about Chebyshev's inequality and compute the variance of the geometric distribution.

The variance gives a powerful way to measure the probability that a random variable deviates from its expectation by a lot. As we have seen, $\mathbb{E}[X]$ does not tell us anything about how far X can be from its expectation. However, we do have the following simple inequality (we proved it in the last lecture):

Fact 1 (Markov's inequality)**.** *If X is a non-negative random variable, then*

$$p(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}.$$

Applying Markov's inequality to the variance gives us Chebyshev's inequality:

Fact 2 (Chebyshev's inequality)**.** *If $\alpha \geq 0$,*

$$p(|X - \mathbb{E}[X]| \geq \alpha) \leq \frac{\text{Var}[X]}{\alpha^2}.$$

Proof.

$$\begin{aligned} p(|X - \mathbb{E}[X]| \geq \alpha) &= p((X - \mathbb{E}[X])^2 \geq \alpha^2) \\ &\leq \frac{\mathbb{E}\left[(X - \mathbb{E}[X])^2\right]}{\alpha^2} \\ &= \frac{\text{Var}[X]}{\alpha^2}. \end{aligned}$$

by Markov's inequality applied to the non-negative random variable $(X - \mathbb{E}[X])^2$.

□

Let us apply Markov and Chebyshev's inequality to the geometric distribution.

Example: Geometric Distribution

Suppose we repeatedly toss a coin until we see heads. Suppose the probability of heads in each coin toss is p. Let X be the number of coin tosses.

We saw in the last lecture that

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} p \cdot (1-p)^{i-1} i = 1/p.$$

To calculate the variance,

$$\mathbb{E}\left[X^2\right] = \sum_{i=1}^{\infty} p \cdot (1-p)^{i-1} i^2 = p \sum_{i=1}^{\infty} (1-p)^{i-1} i^2.$$

To calculate this, we use Taylor series. We have the identity

$$\frac{1}{1-x} = 1 + x + x^2 + \dots.$$

This already shows that

$$\mathbb{E}\left[X\right] = p + p \cdot 2(1-p) + p \cdot 3(1-p)^2 + \cdots = \frac{p}{1-(1-p)}.$$

Taking the derivative gives

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \dots.$$

This already shows that

$$\mathbb{E}\left[X\right] = p + p \cdot 2(1-p) + p \cdot 3(1-p)^2 + \cdots = \frac{p}{(1-(1-p))^2} = \frac{1}{p}.$$

Multiplying the Taylor series by x gives

$$\frac{x}{(1-x)^2} = x + 2x^2 + 3x^3 + 4x^4 + \dots.$$

Then taking the derivative again, we get

$$\frac{1}{(1-x)^2} + \frac{2x}{(1-x)^3} = 1 + 2^2 \cdot x + 3^2 \cdot x^2 + 4^2 \cdot x^3 + \dots,$$

which gives

$$\frac{1+x}{(1-x)^3} = 1 + 2^2 \cdot x + 3^2 \cdot x^2 + 4^2 \cdot x^3 + \dots,$$

so:

$$\mathbb{E}\left[X^2\right] = p \sum_{i=1}^{\infty} (1-p)^{i-1} i^2 = p \cdot \frac{1+1-p}{(1-(1-p))^3} = \frac{2-p}{p^2}.$$

So the variance is:

$$\mathrm{Var}\left[X\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Now, let us estimate the probability that we have to wait twice as long to see the first heads. Markov's inequality gives:

$$p(X \geq 2/p) \leq \frac{\mathbb{E}\left[X\right]}{2/p} \leq 1/2.$$

Chebyshev's inequality gives

$$p(X \geq 2/p) = p(|X - \mathbb{E}[X]| \geq 1/p) \leq \frac{\text{Var}[X]}{1/p^2} = \frac{(1-p)/p^2}{1/p^2} = 1 - p.$$

This is better than Markov's inequality when $p > 1/2$, but worse otherwise.

You can also calculate directly that $p(X \geq 2/p) \leq (1-p)^{2/p-1}$.

What if we toss the coin until we see n heads? Let X denote the number of coin tosses. Let X_1 be the the number of coin tosses to see the first heads, let X_2 denote the number of additional coin tosses to see the second heads and so on. Then we see that

$$X = X_1 + X_2 + \ldots + X_n.$$

Moreover, $X_1, X_2, \ldots, X_n$ are mutually independent. We have

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \ldots + \mathbb{E}[X_n] = n/p,$$

by linearity of expectation. Since the variables are independent, we have

$$\text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_n] = \frac{1-p}{p^2} \cdot n.$$

Let us again estimate the probability that the number of tosses is twice as many as we expect. Markov's inequality gives:

$$p(X \geq 2n/p) \leq \frac{n/p}{2n/p} = \frac{1}{2}.$$

Chebyshev's inequality gives

$$p(X \geq 2n/p) = p(|X - \mathbb{E}[X]| \geq n/p) \leq \frac{\text{Var}[X]}{(n/p)^2} = \frac{(1-p)/p^2 \cdot n}{n^2/p^2} = \frac{1-p}{n}.$$

As n gets larger, Chebyshev's inequality gives a much stronger bound.