Lecture 15: Using Markov and Chebyshev's Inequalities

Anup Rao

February 9, 2018

We continue our discussion of Variance, and talk about using it to prove concentration bounds.

In this lecture, we continue our discussion of using the variance of random variables to prove concentration bounds.

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 the Taylor series approach we discussed last time. We know the identity

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

Taking the derivative gives

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

Multiplying 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:

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

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

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

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

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

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

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

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

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

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