Proof of Markov and Chebyshev's inequalities

Jacob L. Fine

April 29th, 2024

In statistical inference, we often wish to compute an upper bound on the probability that a random variable, which may be a parameter of interest, deviates from some positive value k.

Two inequalities are often used for these questions: Markov's inequality and Chebyshev's inequality. We will derive Markov's inequality and show how we can use it to obtain Chebyshev's inequality.

Markov's inequality states that for some positive value k when θ is non-negative, if the expected value $E(\theta)$ is defined, then

$$P(\theta > k) \leq E(\theta)/k$$

To show why this is true we may write the $E(\theta)$ as

$$E(\theta) = \int_0^\infty \theta f(\theta) d\theta$$

In the above, since θ is non-negative, the lower bound on the integral is zero.

We can rewrite the integral on the LHS to specify some positive value k we are interested in as

$$E(\theta) = \int_0^k \theta f(\theta) d\theta + \int_k^\infty \theta f(\theta) d\theta$$

Since every term in the above is non-negative, the sum of the integrals must be larger than either individual integral, which implies

$$E(\theta) = \int_0^k \theta f(\theta) d\theta + \int_k^\infty \theta f(\theta) d\theta \geq \int_k^\infty \theta f(\theta) d\theta$$

And since k is positive and θ is non-negative, integrating the product $\theta f(\theta)$ will yield a value that is at least as large, or larger, than the integral $\int_k^\infty k f(\theta) d\theta$. Therefore we have

$$E(\theta) \geq \int_k^\infty \theta f(\theta) d\theta \geq k \int_k^\infty f(\theta) d\theta$$

And since $\int_k^\infty f(\theta)d\theta$ is just the probability that θ is larger than k, we can write

$$E(\theta) \geq kP(\theta > k)$$
$$\implies P(\theta > k) \leq E(\theta)/k$$

which is the expression of Markov's inequality. We may now use this to derive Chebyshev's inequality, which is

$$P(|\theta - E(\theta)| > k) \leq Var(\theta)/k^2$$

To show this, we may treat $|\theta - E(\theta)|$ as a non-negative random variable and substitute it into Markov's inequality

$$\implies P(|\theta - E(\theta)| > k) \leq E(|\theta - E(\theta)|)/k$$

We may square both sides on the inequality inside the probability, as well as the term of the RHS $|\theta - E(\theta)|/k$, to obtain

$$\implies P([\theta - E(\theta)]^2 > k^2) \leq E[\theta - E(\theta)]^2/k^2$$

And since the definition of variance is

$$E([\theta - E(\theta)]^2) = Var(\theta)$$

We may write

$$\implies P(|\theta - E(\theta)| > k) \leq Var(\theta)/k^2$$

which is the desired result.

Now we will use Chebyshev's inequality in a sample problem. Let X be the number of successful independent drug treatments, with the $Binom(n, p)$ where $p = 0.9$. Suppose we are interested in obtain an upper bound on the probablity that 95% or more of the drug treatments are successful, i.e.,

$$P(X > 0.95n)$$

To make this into a form workable with Chebyshev's inequality, we can subtract the mean $0.90n$ from both sides such that $E(X) = np$ appears on the RHS. Therefore

$$P(X - 0.9n > 0.05n)$$

We can now use Chebyshev's inequality to obtain an upper bound, where $k = 0.05n$ and $Var(X) = np(1 - p)$, since the distribution is binomial. Therefore

$$P(X - 0.9n > 0.05n) \leq 0.9n(1 - 0.9)/(0.05n)^2$$
$$\implies P(X - 0.9n > 0.05n) \leq 36/n$$

The tightness of the upper bound improves when n increases.