On Chebyshev's and Kolmogorov's Inequalities

September 29, 2024

Let X be a random variable (RV) with expectation $\mathrm{E}(X)$ and variance $\mathrm{var}(X)$, and let $\lambda > 0$. The Chebyshev's inequality says

$$\Pr(|X - \mathrm{E}(X)| \geq \lambda) \leq \frac{\mathrm{var}(X)}{\lambda^2}$$

Let $X_1, \dots, X_n$ be IID RV's with expectation μ and variance σ^2. For $k = 1, \dots, n$, let $S_k = X_1 + \cdots + X_k$. Then $\mathrm{E}(S_k) = k\mu$, $\mathrm{var}(S_k) = k\sigma^2$, and we get

$$\Pr(|S_k - k\mu| \geq \lambda) \leq \frac{k\sigma^2}{\lambda^2}$$

The upper bounds reflect the fact that $\{|S_k - k\mu| \geq \lambda\}$ is more likely for a larger k.

Consider the special case of tossing a fair coin, i.e., $X_i \sim \mathrm{Bernoulli}(0.5)$, with $\mathrm{E}(X_i) = 1/2$, $\mathrm{var}(X_i) = 1/4$. Then S_k is the total number of heads in k tosses. For $n = 100$ and $\lambda = 10$,

$$\Pr(|S_{100} - 50| \geq 10) \leq \frac{1}{4}$$

In other words, there is at most a 1/4 chance that there are ≤ 40 or ≥ 60 heads. The actual chance is 0.04 to two decimal places, by summing up binomial probabilities.

Imagine repeating 100 tosses of a fair coin many (say, 10,000) times, and storing the results (1 for head, 0 for tail) in a matrix of 100 columns and many rows. Then each row sum is the number of heads observed in a run of the experiment. Let p be the proportion of rows with sum ≤ 40 or ≥ 60. Chebyshev's inequality says $p \leq 1/4$, up to random fluctuations.

Now colour a row red if any of the following occurs

- The sum of all 100 numbers is ≥ 10 away from 50.
- The sum of the first 99 numbers is ≥ 10 away from 49.5.
- ...
- The sum of the first k numbers is ≥ 10 away from $k/2$.
- ...

- The sum of the first 2 numbers is ≥ 10 away from 1.
- The first number is ≥ 10 away from 1/2.

Obviously, the last 19 conditions are impossible. Still, there are 80 others to make the red rows more numerous than those with sum ≤ 40 or ≥ 60. Therefore, $p \leq p_r$, the proportion of red rows. Yet, $p_r \leq 1/4$, roughly. This is a consequence of

Kolmogorov's inequality
Let $X_1, \ldots, X_n$ be IID RV's with expectation μ and variance σ^2. Let $S_k = X_1 + \cdots + X_k$, $k = 1, \ldots, n$. Then for any $\lambda > 0$,

$$\Pr\left(\max_{1 \leq k \leq n} |S_k - k\mu| \geq \lambda\right) \leq \frac{n\sigma^2}{\lambda^2}$$

Notice that $\max_{1 \leq k \leq n} |S_k - k\mu| \geq \lambda$ exactly when any of the absolute differences is at least λ, which is like the definition of a red row. It is quite remarkable that Kolmogorov slipped a larger event into Chebyshev's inequality for S_n.

A computer simulation of coin tossing shows that

$$\Pr\left(\max_{1 \leq k \leq 100} |S_k - k/2| \geq 10\right) \approx 0.09$$

which is more than double of $\Pr(|S_{100} - 50| \geq 10)$, but still quite a bit smaller than 1/4. Kolmogorov's inequality is also quite generous.

This is dedicated to my mother, who guided me to mathematics and music.