(Weak) Law of Large Numbers

One of the important ideas in probability is that of interpreting the results of a large number of independent experiments.

As a self-referential example, we can think of the task of finding mistakes in Mathematics! Each (serious) student of Mathematics conducts an experiment of trying to find a mistake in a "standard" result. Assuming that each student does this independently (and does not just believe the teacher or her/his friend!), we have a sequence of independent experiments! So the question we can ask is: "What is the confidence that we have that the result is correct given that no one has found a mistake so far?"

Another application is the famous quote about Free and Open source Software: Given enough eyes all bugs are shallow!

Momentary Inequalities

The following inequalities are *not* momentary but are well-established! However, they are inequalities involving moments of random variables. We will give proofs for discrete random variables, the proofs for a general random variable follow from the application of limiting techniques.

Given a real-valued random variable X with a finite value for $E(|X|^k)$. For any $c > 0$ we have:

$$E(|X|^k/c^k) = \sum_{a \in D} P(X = a)|a|^k/c^k \geq \sum_{\substack{a \in D \\ |a| \geq c}} P(X = a) = P(|X| \geq c)$$

From this we see that

$$P(|X| \geq c) \leq \frac{E(|X|^k)}{c^k}$$

For $k = 2$ this is called Chebyschev's Inequality and for $k = 1$ this is called Markov's Inequality.

Weak Law of Large Numbers

A sequence of independent experiments is mathematically represented by a sequence of independent random variables X_i for $i = 1, \dots, .$ We assume that $E(X_i) = m_i$ and $\sigma^2(X_i) \leq M$ for some fixed $M > 0$. Let $Y_n = (X_1 + \cdots + X_n)/n$ be the random variable that averages these random variables, then

$$E(Y_n) = \frac{E(X_1) + \cdots + E(X_n)}{n} = \frac{m_1 + \cdots + m_n}{n}$$

The weak Law of Large Numbers says that for any $c > 0$,

$$P(|Y_n - E(Y_n)| \geq c) \leq \frac{M}{c^2 n}$$

In particular, this goes to 0 as n goes to infinity.

By the independence of the random variables X_i, we have

$$\sigma^2(X_1 + \cdots + X_n) = \sigma^2(X_1) + \cdots + \sigma^2(X_n) \leq nM$$

It follows that

$$\sigma^2(Y_n - E(Y_n)) = \sigma^2(Y_n) = (1/n^2)\sigma^2(X_1 + \cdots + X_n) \leq \frac{M}{n}$$

The result follows from an application of Chebyschev's inequality to the random variable $Y_n - E(Y_n)$.

Measuring a Physical Quantity

If the sequence of experiments are many different experiments to measure the same Physical Quantity (here "Physical" includes Chemical and Biological!), then we have $E(X_i) = m_i = m$ for all the experiments. It follows that $E(Y_n) = m$ as well.

We assume that we can control the error in each experiment to ensure that $\sigma^2(X_i) \leq M$ for some fixed constant M. (Typically, M is a measure of control we have over the experiment; smaller M means more control.)

An application of the weak Law of Large Numbers gives us the estimate

$$P(|Y_n - m| > 1/10^k) = \frac{10^{2k}M}{n}$$

Thus, we can get a precise estimate on the number of experiments required to make our confidence in the result as high as we wish.

Frequency and Probability

We apply the weak Law of Large Numbers to the random variables that are a sequence of independent coin flips with probability p of Head. The random variable X_i has probabilities given by $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$. Then $E(X_i) = p$ and $\sigma^2(X_i) = p(1-p)^2 + (1-p)p^2 = p(1-p)$.

As above, we see that $E(Y_n) = p$. On the other hand Y_n counts the *frequency* of heads in the coin flips. The weak Law of Large numbers

$$P(|Y_n - p| > c) \to 0 \text{ as } n \to \infty$$

justifies our intuition that, with high probability, the frequency of occurence of heads is the same as the probability p once we carry out a large number of coin flips.