In this lecture, we develop the weak law of large numbers.

Loosely speaking, the weak law of large numbers says that if we have a sequence of independent random variables with the same distribution, then the average of these random variables, which is called the sample mean, approaches the expected value of the distribution.

In this sense, it reinforces our interpretation of the expected value as some kind of overall average.

The weak law of large numbers is the reason why polling works.

By asking many people about the value of some attribute, and by taking the average of the responses, we can get a good estimate of the average over the entire population.

On the mathematical side, in order to derive the weak law of large numbers, we will first need to develop some inequalities, namely the Markov and Chebyshev inequalities.

Both of them tell us something about tail probabilities.

Suppose that a is a number.

Then it is reasonable to expect that the probability that the random variable exceeds a will be small when a is very large.

But how small?

The Markov and Chebyshev inequalities give us some answers to this question, based only on knowledge of the mean and the variance of the distribution.

Finally, we will have to deal with a technical issue.

The weak law of large numbers talks about the convergence of a random variable to a number.

For this to make sense, we need to define an appropriate notion of convergence.

We will introduce one such notion that goes under the name of convergence in probability.

And we will see that in many respects, it is similar to the common notion of convergence of numbers.