The Law of Large Numbers

If $X_1, X_2, \ldots, X_N$ are pairwise independent, identically distributed random variables with expectation $E(X)$ and variance $V(X)$ then:

the probability that

$$\left| \frac{X_1 + X_2 + \ldots + X_N}{N} - E(X) \right| \leq \varepsilon$$

is greater than $1 - \frac{V(X)}{N\varepsilon^2}$ where $\varepsilon$ is a positive number.

Proof:

We will apply Chebyshev’s Inequality to the random variable

$$\frac{X_1 + X_2 + \ldots + X_N}{N}.$$

First we find the expectation and variance of this random variable using the formulas from my paper on the algebra of expectations.

Since the expectation of a sum of random variables equals the sum of their expectations we get: $E(X_1 + X_2 + \ldots + X_N) = NE(X)$.

Since the expectation of a constant times a random variable is the product of the constant times the expectation of the random variable we get: $E\left( \frac{X_1 + X_2 + \ldots + X_N}{N} \right) = 1/N \times NE(X) = E(X)$. 
Since the variance of the sum of pairwise independent random variables is the sum of their variances we get:

\[ V(X_1 + X_2 + \ldots + X_N) = NV(X) . \]

Since the variance of a constant times a random variable is equal to the constant squared times the variance of the random variable we get:

\[
V \left( \frac{X_1 + X_2 + \ldots + X_N}{N} \right) = \frac{NV(X)}{N^2} = \frac{V(X)}{N} 
\]

Chebyshev’s Inequality says that:

the probability that \(|Y - E(Y)| \leq \varepsilon\) is greater than \(1 - \frac{V(Y)}{\varepsilon^2}\).

So since \(E(\frac{X_1 + X_2 + \ldots + X_N}{N}) = E(X)\) and \(V(\frac{X_1 + X_2 + \ldots + X_N}{N}) = \frac{V(X)}{N}\), we get when substituting \(\frac{X_1 + X_2 + \ldots + X_N}{N}\) for \(Y\) in Chebyshev’s Inequality:

the probability that \(\left| \frac{X_1 + X_2 + \ldots + X_N}{N} - E(X) \right| \leq \varepsilon\) is greater than \(1 - \frac{V(X)}{N\varepsilon^2}\).
Proof of Bernoulli’s Theorem using Chebyshev’s Law of Large Numbers

Let $X_i = 1$ if there is a success on the ith trial
Let $X_i = 0$ if there is a failure on the ith trial

let $p$ be the probability of a success on the ith trial
let $q = 1 - p$ be the probability of a failure on the ith trial

then $E(X_i) = 1 \times p + 0 \times q = p$

and $V(X_i) = (1 - p)^2 p + (0 - p)^2 q = q^2 p + p^2 q = qp(q+p) = pq$

So by Chebyshev’s law of large numbers we have:

the probability that $\left| \frac{X_1 + X_2 + \ldots + X_N}{N} - p \right| \leq \varepsilon$ is greater than $1 - \frac{pq}{N\varepsilon^2}$.

Let $M$ be the actual number of successes in $N$ trials.

$M = X_1 + X_2 + \ldots + X_N$ so we get:

the probability that $\left| \frac{M}{N} - p \right| \leq \varepsilon$ is greater than $1 - \frac{pq}{N\varepsilon^2}$.

Let $\eta$ be a positive number, then if $N$ is sufficiently large, the probability that $\left| \frac{M}{N} - p \right| \leq \varepsilon$
is greater than $1 - \eta$.

This is Bernoulli’s Theorem.
Comment

In Bernoulli’s example, \( p = 3/5 \), \( q = 2/5 \), \( e = 1/50 \), and

\[
\frac{pq}{Ne^2} = 1/1001.
\]

Solving for \( N \) we get \( N = 600,600 \).

Using Bernoulli’s formula, \( NT \) is 25,550 where Bernoulli’s \( NT \) is our \( N \). So Bernoulli’s formula gives a much better value for \( N \).

If we modify Bernoulli’s example so that \( C \) is 9 instead of 1000, then \( NT \) becomes 13,900 and using Chebyshev’s inequality the number of trials becomes 6,000. So when \( C \) is small, Chebyshev’s inequality can give a smaller number of trials than Bernoulli’s formulas.

Daniel Daniels  updated 11/22/2020