The following law of large numbers was discovered by Jacob Bernoulli (1655–1705). Both the statement and the way of its proof adopted today are different from the original.

**Theorem** Let a particular outcome occur with probability $p$ as a result of a certain experiment. Let the experiment be repeated independently over and over again, and let $\hat{m} = \hat{m}(n)$ be the observed frequency of the outcome, i.e., the number of times that the outcome occurs in $n$ trials. Then for any positive $\varepsilon$, no matter how small,

$$\text{the probability that } \frac{\hat{m}}{n} \text{ and } p \text{ differ by less than } \varepsilon$$

approaches 1, as $n$ increases without bound.

**Proof** Let $X_i$ be the 1/0 indicator of whether the outcome occurs in the $i$th trial. Then $X_1 + \ldots + X_n = \hat{m}$.

We have $\mathbb{E}[X_i] = p$ and $\text{Var}(X_i) = p - p^2$. So $\mathbb{E}[\hat{m}/n] = p$ and $\text{Var}(\hat{m}/n) = \frac{p(1-p)}{n}$.

Fix any $\varepsilon > 0$. By Chebyshev’s inequality,

$$P\left(\left|\frac{\hat{m}}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{p(1-p)}{n \varepsilon^2}.$$  

But the right side of the inequality increases to 1, as $n \to \infty$.  

For a fixed $\varepsilon$, this is how $P\left(\left|\frac{\hat{m}}{n} - p\right| < \varepsilon\right)$ may look like as a function of $n$.

Thus, after sufficiently many trials, the observed relative frequency of an outcome will, with a high degree of certainty, stay within a given $\varepsilon$ of the outcome probability. As the number of trials increases without bound, the degree of certainty increases to 1. One says that, as $n \to \infty$, the relative frequency $\hat{m}/n$ converges to $p$ in probability.

---

1 J. Bernoulli, *Ars Conjectandi*, Basel, 1713