STRONG AND WEAK LAW OF LARGE NUMBERS

In probability theory, the law of large numbers (LLN) is a theorem that describes the result of performing the same experiment a large number of times. According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.

The sequence of variates $X_i$ with corresponding means $\mu_i$ obeys the strong law of large numbers if, to every pair $\epsilon, \delta > 0$, there corresponds an $N$ such that there is probability $1 - \delta$ or better that for every $r > 0$, all $r + 1$ inequalities

$$\frac{|S_n - m_n|}{n} < \epsilon$$

for $n = N, N + 1, ..., N + r$ will be satisfied, where

$$S_n \equiv \sum_{i=1}^{n} X_i$$

$$m_n \equiv \langle S_n \rangle = \mu_1 + ... + \mu_n$$

(Feller 1968). Kolmogorov established that the convergence of the sequence

$$\sum \frac{\sigma_k^2}{k^2},$$

sometimes called the Kolmogorov criterion, is a sufficient condition for the strong law of large numbers to apply to the sequence of mutually independent random variables $X_k$ with variances $\sigma_k$. 
WEAK LAW OF LARGE NUMBERS

The weak law of large numbers (cf. the strong law of large numbers) is a result in probability theory also known as Bernoulli’s theorem. Let \( X_1, ..., X_n \) be a sequence of independent and identically distributed random variables, each having a mean \( \langle X_i \rangle = \mu \) and standard deviation \( \sigma \).

Define a new variable

\[
X \equiv \frac{X_1 + \ldots + X_n}{n}. \tag{1}
\]

Then, as \( n \to \infty \), the sample mean \( \langle X \rangle \) equals the population mean \( \mu \) of each variable.

\[
\langle X \rangle = \frac{X_1 + \ldots + X_n}{n} \tag{2}
\]

\[
= \frac{1}{n} (\langle X_1 \rangle + \ldots + \langle X_n \rangle) \tag{3}
\]

\[
= \frac{n \mu}{n} \tag{4}
\]

\[
= \mu. \tag{5}
\]

In addition,

\[
\text{var} (X) = \text{var} \left( \frac{X_1 + \ldots + X_n}{n} \right) \tag{6}
\]

\[
= \text{var} \left( \frac{X_1}{n} \right) + \ldots + \text{var} \left( \frac{X_n}{n} \right) \tag{7}
\]

\[
= \frac{\sigma^2}{n^2} + \ldots + \frac{\sigma^2}{n^2} \tag{8}
\]

\[
= \frac{\sigma^2}{n}. \tag{9}
\]
Therefore, by the Chebyshev inequality,

\[ P (|X - \mu| \geq \epsilon) \leq \frac{\text{var} (X)}{\epsilon^2} = \frac{\sigma^2}{n \epsilon^2}. \]  

(10)

As \( n \to \infty \), it then follows that

\[ \lim_{n \to \infty} P (|X - \mu| \geq \epsilon) = 0. \]  

(11)

Stated another way, the probability that the average \( \frac{(X_1 + \ldots + X_n)}{n} - \mu \) approaches \( \epsilon \) for an arbitrary positive quantity approaches 1 as \( n \to \infty \).

**MARKOV’S INEQUALITY**

If \( x \) takes only nonnegative values, then

\[ P (x \geq a) \leq \frac{\langle x \rangle}{a}. \]  

(1)

To prove the theorem, write

\[ \langle x \rangle = \int_0^\infty x P (x) \, dx \]  

(2)

\[ = \int_0^a x P (x) \, dx + \int_a^\infty x P (x) \, dx. \]  

(3)

Since \( P (x) \) is a probability density, it must be \( \geq 0 \). We have stipulated that \( x \geq 0 \), so

\[ \langle x \rangle = \int_0^a x P (x) \, dx + \int_a^\infty x P (x) \, dx. \]  

(4)
\[ \begin{align*}
&\forall x \int_{x}^{\infty} P(x) \, dx \\
&\forall a \int_{a}^{\infty} a \, P(x) \, dx \\
&= a \int_{a}^{\infty} P(x) \, dx \\
&= a P(x \geq a).
\end{align*} \]

CHEBYSHEV INEQUALITY

Apply Markov's inequality with \( a \equiv k^2 \) to obtain

\[ P \left[ (x - \mu)^2 \geq k^2 \right] \leq \frac{(x - \mu)^2}{k^2} = \frac{\sigma^2}{k^2}. \] (1)

Therefore, if a random variable \( x \) has a finite mean \( \mu \) and finite variance \( \sigma^2 \), then for all \( k > 0 \),

\[ P \left( |x - \mu| \geq k \right) \leq \frac{\sigma^2}{k^2} \] (2)

\[ P \left( |x - \mu| \geq k \sigma \right) \leq \frac{1}{k^2}. \] (3)