

Weak Law of Large Numbers

Willard Thór Rafnsson

25th January 2009

Definition 1 (Expected Value (mean))

Let X be a discrete random variable with image $\mathcal{X} = \{x_1, x_2, \dots\}$ and probability mass function $p : \mathcal{X} \rightarrow [0, 1]$. The *expected value* of X is defined as

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} x_k p(x_k). \quad \square$$

When the random variable is clear from the context, we denote its mean as μ .

Definition 2 (Variance)

Let X be a random variable with mean μ . The variance of x is defined as

$$\text{Var}[X] = \mathbb{E}[(X - \mu)^2]. \quad \square$$

When the random variable is clear from the context, we denote its variance as σ^2 , where $\sigma \geq 0$ is called the (standard) deviation.

Theorem 1 (Chebyshev's Inequality)

Let x be a random variable with image $\{x_1, x_2, \dots\}$, mean μ and variance σ^2 . For all constants $c > 0$, it holds that

$$\Pr(|X - \mu| \geq c\sigma) \leq \frac{1}{c^2}. \quad (1)$$

PROOF

We prove this theorem for discrete random variables only¹. Let $p : \mathcal{X} \rightarrow [0, 1]$ be a probability mass function for X . Let $c > 0$ be given and let $B = \{x_i \in \mathcal{X} : |x_i - \mu| \geq c\sigma\}$. By definition of variance and mean, we have

$$\sigma^2 = \mathbb{E}[(X - \mu)^2] = \sum_{i=1}^{\infty} (x_i - \mu)^2 p(x_i).$$

Since all of B is in the sum, we have

$$\sigma^2 \geq \sum_B (x_i - \mu)^2 p(x_i) \geq c^2 \sigma^2 \sum_B p(x_i) = c^2 \sigma^2 \Pr(X \in B),$$

¹A proof for the continuous case can be found in [Olo05, Proposition 2.4.7]

where we used the observation that $(x_i - \mu)^2 \geq c^2 \sigma^2$ by definition of B . Chebyshev's inequality follows by division with $\sigma^2 c^2$. ■

Chebyshev's inequality thus gives us a lower bound on the probability of X obtaining a value within a given distance $c\sigma$ from the expected value μ , given the variance σ^2 :

$$\Pr(|X - \mu| < c\sigma) > 1 - \frac{1}{c^2}.$$

The *law of large numbers* expresses that the empiric mean for an infinite sequence of random variables converges towards the actual mean. The following is a *weakened* version of the law of large numbers, where we assume that the variance of the infinite sequence of random variables is finite.

Theorem 2 (Weak Law of Large Numbers)

Let X_1, X_2, \dots be a sequence of iid² random variables with $E[X_i] = \mu$, $\text{Var}[X_i] = \sigma^2$ for $i = 1, \dots, n$, and let $S_n = X_1 + \dots + X_n$. It then holds that

$$\forall \delta > 0. \lim_{n \rightarrow \infty} \Pr\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) = 0,$$

where $\frac{S_n}{n}$ is the sample mean.

PROOF

Assume that the sequence of random variables has a finite length n . Since $E[\cdot]$ is linear³, we have

$$E\left[\frac{S_n}{n}\right] = \frac{1}{n} (E[X_1] + \dots + E[X_n]) = \frac{n\mu}{n} = \mu.$$

From the addition- and multiplication rules of $\text{Var}[\cdot]$ ⁴, we have

$$\begin{aligned} \text{Var}\left[\frac{S_n}{n}\right] &= \frac{1}{n^2} (\text{Var}[X_1] + \dots + \text{Var}[X_n]) \\ &= \frac{1}{n^2} \text{Var}[X_1 + \dots + X_n] \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}, \end{aligned}$$

Thus the deviation for $\frac{S_n}{n}$ is $\frac{\sigma}{\sqrt{n}}$. By insertion into Chebyshev's inequality, (1), with constant $c = \delta \frac{\sqrt{n}}{\sigma}$ for $\delta > 0$, we get

$$\begin{aligned} \Pr\left(\left|\frac{S_n}{n} - \mu\right| \geq \left(\delta \frac{\sqrt{n}}{\sigma}\right) \frac{\sigma}{\sqrt{n}}\right) &= \Pr\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) \\ &\leq \frac{\sigma^2}{\delta^2 n}. \end{aligned}$$

The result follows by taking the limit for $n \rightarrow \infty$ in the above inequality. ■

²Independent, Identically Distributed

³This is proven in [Olo05, Proposition 3.6.6]

⁴[Olo05, Proposition 3.6.7 b)] and [Olo05, Proposition 3.6.4]

References

- [Olo05] Peter Olofsson. *Probability, Statistics, and Stochastic Processes*. John Wiley & Sons, Inc., 2005.