§ 1.7 Law of Large Numbers

A dice is rolled many times. The number of times the point $i (i = 1, 2, \cdots, 6)$ appears relative to the total number of rolls (which is the **frequency** of i points) is about $1/6$. It is called the law of large number, and everyone must have experienced it. The objective of this section is to represent this empirical fact in a mathematically correct manner, and to prove it.

Since the probability space (Ω_0, P_0) of the trial of rolling a dice is given by

$$\Omega_0 = \{1, 2, \cdots, 6\} \qquad P_0 i = 1/6\,, \quad i = 1, 2, \cdots, 6\,.$$

The probability space (Ω, P) of the trial of rolling a dice n times is given by

$$\Omega = \Omega_0^n\,, \qquad P = P_0^n\,.$$

The point that appears for the k-th roll is denoted by $X_k = X_k(\omega)$ and

$$X_k(\omega) = \pi_k(\omega) \qquad ((\pi_k \text{ is the } k \text{ image})\,, \qquad k = 1, 2, \cdots, n\,.$$

Clearly, $X_1, X_2, \cdots, X_n$ are independent probability variables on (Ω, P),

$$P\{X_k(\omega) = i\} = 1/6\,, \qquad k = 1, 2, \cdots, n\,.$$

Now, using $1_i : \Omega \longrightarrow \{0, 1\}\,, 1_i(i) = 1\,, 1_i(j) = 0\,(j \neq i)$, the frequency R_i of i point up to n times is given by

$$R_i = R_i(\omega) = \frac{1}{n}\sum_{k=1}^{n} 1_i\big(X_k(\omega)\big)\,.$$

Hence, R_i is a probability variable on (Ω, P). Though R_i is related to n, to make the symbol less cumbersome, it is omitted. The law of large numbers assets that when n is sufficiently large,

$$R_i(\omega) \doteqdot 1/6\,, \qquad i = 1, 2, \cdots, 6\,.$$

The foremost question is the meaning of this approximation. As the left hand side is a function of ω, it can be considered to mean

$$\max_{\omega} \left| R_i(\omega) - \frac{1}{6} \right| < \varepsilon \qquad (\,\varepsilon \text{ is a very small positive number})\,.$$

When the number of times n is made larger, it is possible that the point i appears consecutively n times, and since $R_i(\omega) = 1$, the above interpretation is inappropriate. That said, the probability of the extreme case of $R_i(\omega) = 1$ when n is large, is extremely small. Thus, the interpretation of the approximation is taken as

$$\text{“}P\left\{\left|R_i - \frac{1}{6}\right| > \varepsilon\right\} \text{ is extremely small.”}$$

This is correct, and indeed can be proven. More precisely, it is as follows:

Theorem 1.26 (**Law of Large Numbers**) Suppose with respect to $\varepsilon > 0$, $n_0(\varepsilon)$ is sufficiently large. Then for all $n > n_0(\varepsilon)$,

$$P\left\{\left|R_i - \frac{1}{6}\right| > \varepsilon\right\} < \varepsilon\,, \qquad i = 1, 2, \cdots, 6\,.$$

Proof Fix i, and let $e_k(\omega) = 1_i\big(X_k(\omega)\big)$. Since $X_1, X_2, \cdots, X_n$ are independent, $e_1, e_2, \cdots, e_n$ are also independent. From the additive property of variance,

$$\begin{aligned}
&V(R_i) = V\left(\frac{e_1 + e_2 + \cdots + e_n}{n}\right) = \frac{1}{n^2}V(e_1 + e_2 + \cdots + e_n) = \frac{1}{n^2}\big(V(e_1) + V(e_2) + \cdots + V(e_n)\big)\,,\\
&E(e_k) = 1P\{X_k = i\} + 0P\{X_k \neq i\} = \frac{1}{6}\,,\\
&E(R_i) = E\left(\frac{e_1 + e_2 + \cdots + e_n}{n}\right) = E\left(\frac{E(e_1) + E(e_2) + \cdots + E(e_n)}{n}\right) = \frac{1}{6}\,,\\
&V(e_i) = E\left(\left(e_i - \frac{1}{6}\right)^2\right) = \left(\frac{5}{6}\right)^2 P\{X_k = i\} + \left(\frac{1}{6}\right)^2 P\{X_k \neq i\} = \frac{5}{36}\,,\\
&V(R_i) = \frac{5}{36n}\,, \qquad \sigma(R_i) = \frac{\sqrt{5}}{6\sqrt{n}}\,.
\end{aligned}$$

From the Čebysěv's inequality,

$$P\left\{\left|R_i - \frac{1}{6}\right| > a\frac{\sqrt{5}}{6\sqrt{n}}\right\} \leqq \frac{1}{a^2}\,.$$

Take $a = a(\varepsilon)$ to be sufficiently large and the right hand side of the above inequality becomes smaller than ε. With respect to this a, $n_0 = n_0(\varepsilon)$ is taken to be sufficiently large, and so long as $n > n_0$,

$$a\frac{\sqrt{5}}{6\sqrt{n}} < a\frac{\sqrt{5}}{6\sqrt{n_0}} < \varepsilon\,.$$

Hence, so long as $n > n_0(\varepsilon)$,

$$P\left\{\left|R_i - \frac{1}{6}\right| > \varepsilon\right\} < \varepsilon\,. \qquad \blacksquare$$

Exercise 1.7 With respect to a general trial that is repeated, present mathematically the law of large numbers and prove it.

Christopher Ting
Translation Version 1.0
Date: 10 May, 2017