1. Convergence in probability

First we define a mode of convergence for sequences of random variables.

Recall that a random variable is a *function* that maps the elements of a sample space S into the real numbers. This means that a sequence of random variables is in fact a sequence of functions, a concept we have encountered before.

For example, the sequence of partial sums of a power series is a sequence of functions. If the power series converges, we then have a sequence of functions that converges to a limiting function.

Definition (convergence in probability). *Let $\{Z_n\}$ be a sequence of random variables,*

$$\{Z_n\} = Z_1, Z_2, Z_3, \ldots$$

We say that $\{Z_n\}$ converges to the random variable Z in probability and write

$$Z_n \xrightarrow{P} Z$$

if, for every $\epsilon > 0$,

$$P(|Z_n - Z| \geq \epsilon) \to 0 \quad \textit{as} \quad n \to \infty.$$

In effect, the definition says that the probability that Z_n and Z differ by any given amount (ϵ) will be negligible if n is large enough.

2. Bernoulli's (weak) Law of Large Numbers

Theorem (Bernoulli). *If Y_n is a sequence of random variables having the binomial distribution*

$$Y_n = \mathcal{B}(n, p), \quad n = 1, 2, 3, \ldots$$

then

$$\frac{Y_n}{n} \xrightarrow{P} p$$

That is, as n becomes large, the probability that Y_n/n differs from p by any given amount tends to zero.

Note that our definition of convergence in probability requires that the symbol p following $\xrightarrow{P}$ represent a random variable. So, strictly speaking, p should not be interpreted as a constant, but rather as the random variable that takes the value p with probability 1, or, equivalently, the random variable that maps every element of the sample space to the constant p. That said, the random variable p behaves exactly like the constant p.

Proof. From Theorem 3.7, we know that since Y_n is distributed as binomial (n, p), then

$$E(Y_n) = np \quad \text{and} \quad V(Y_n) = np(1-p)$$

From Exercise 3.33, we have that if Y is a random variable and a and b are constants, then

$$E(aY + b) = aE(Y) + b$$

so that if we let $a = 1/n$ and $b = -p$, we have

$$E\left(\frac{Y_n}{n} - p\right) = \frac{1}{n}E(Y_n) - p = \frac{1}{n} \cdot np - p = p - p = 0$$

Also from 3.33, we know that

$$V(aY + b) = a^2 V(Y)$$

so

$$V\left(\frac{Y_n}{n} - p\right) = \frac{1}{n^2}V(Y_n) = \frac{n \cdot p(1-p)}{n^2} = \frac{p \cdot (1-p)}{n}$$

Now define

$$S_n = \left(\frac{Y_n}{n} - p\right), \quad n = 1, 2, 3, \ldots$$

Since $E(S_n) = 0$,

$$V(S_n) = \frac{p \cdot (1-p)}{n} = E(S_n^2) - [E(S_n)]^2 = E(S_n^2)$$

Now suppose $\epsilon > 0$ is given. Then by Chebychev's inequality,

$$P(|S_n| \geq \epsilon) \leq \frac{E(S_n^2)}{\epsilon^2}$$

and by substitution we get

$$P(|S_n| \geq \epsilon) \leq \frac{p \cdot (1-p)}{n\epsilon^2} \to 0$$

as $n \to \infty$ because p and ϵ are constants. $\square$