Law of Large Numbers

Suppose we have $x_i : \Omega \to \mathbb{R}^d \quad i = 1, ..., \infty$ variables.

Definition (*Independence*)

$$P(x_{i1} \,\epsilon\, A_1, ..., x_{ik} \,\epsilon\, A_k) = P(x_{i1} \,\epsilon\, A_1) \cdots P(x_{ik} \,\epsilon\, A_k)$$

Definition (*Identically distributed*)

$$law(x_i) = law(x_j)$$

Theorem (*Law of Large Numbers*)
If $x_i : \Omega \to \mathbb{R}^d \quad i = 1, ..., \infty$ are independent and identically distributed (iid) then,

$$\frac{x_1 + ... + x_n}{n} \cong E(X) \;\; for\ n\ large\ enough$$

* specify strong law vs. weak

Additionally, for any two points , w and a,

$$\frac{x_1(w) + ... + x_n(w)}{n} - \frac{x_1(a) + ... + x_n(a)}{n} \to 0 \quad when\ n \to \infty$$

only true when using strong law

except when $a, w \,\epsilon\, N\ and\ P(N) = 0$

Essentially, for n large enough if we know the expectation at one point we know the expectation at any other point.

The Law of Large numbers is often used in conjunction with the Central Limit theorem.

Theorem *(Central Limit Theorem)* (CLT) If $x_1, ..., x_n$ is a random sample from a distribution with mean μ and variance $\sigma^2 < \infty$ then the limiting distribution of

$$Z_n = \frac{\sum_{i=1}^n X_i \;\; - n\mu}{\sqrt{n}\sigma}$$

is the standard normal, $Z_n \xrightarrow{d} Z \sim N(0, 1)$ as $n \to \infty$.

The key idea behind the CLT is that is can be used to approximate a distribution in cases where the exact distribution is unknown or intractable.

Remarks

- n = 30 is sufficiently large for the approximations using the CLT.
- The average of the sample means and standard deviations will equal the population mean and standard deviation.