# Law of Large Numbers (Section 5.2) 

Let $X_1, X_2, \ldots$ be a sequence of independent r.v.'s with $E(X_i)=\mu$ and $Var(X_i)=\sigma^2$.

Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ for each $n \in \mathbb{N}$ ← Use Y's here instead of X's

For any $\varepsilon > 0$,

$$P(|\bar{X}_n - \mu| > \varepsilon) \to 0 \text{ as } n \to \infty$$

Intuition: If $n$ is large enough, sample mean is close to $\mu$.
Still true if $Y_1, Y_2, \ldots$ are "not too dependent"

---

Application: Suppose we want to estimate

$$E[g(X)] = \int g(x) \cdot f_X(x)\, dx$$

Example 1: If $g(x)=x$, $E[g(X)] = E(X) = \int x f_X(x)\, dx$

Example 2: If $g(x) = \mathbb{I}_{[a,b]}(x) = \begin{cases} 1 & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$ then

$$E[g(x)] = \int \mathbb{I}_{[a,b]}(x) f_X(x)\, dx$$

$$= \int_{-\infty}^{a} \mathbb{I}_{[a,b]}(x) f_X(x)\, dx + \int_a^b \mathbb{I}_{[a,b]}(x) f_X(x)\, dx + \int_b^{\infty} \mathbb{I}_{[a,b]}(x) f_X(x)\, dx$$

$$= \int_{-\infty}^{a} 0 f_X(x)\, dx + \int_a^b 1 \cdot f_X(x)\, dx + \int_b^{\infty} 0 f_X(x)\, dx$$

$$= P(X \in [a,b])$$

Algorithm:

1) Draw $X_1, \ldots, X_n \sim f_X(x)$ (same distribution as X)
2) Set $Y_1 = g(X_1), \ldots, Y_n = g(X_n)$
3) Calculate $\frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$

$$E(Y_i) = E[g(X_i)] = E[g(X)]$$

By the law of large numbers, if n is large enough,

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i) \approx E[g(X)]$$

Example 1: We can estimate $E(X)$ by $\frac{1}{n}\sum_{i=1}^{n} X_i$

Example 2: We can estimate $P(X \in [a,b])$ by

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{[a,b]}(X_i) = \frac{1}{n}(\# \text{ of } X_i \text{ in the interval } [a,b])$$

$=$ proportion of sampled X's in the interval $[a,b]$.

Note: If $W, X$ are jointly distributed r.v.'s and we have samples $(w_1, x_1), (w_2, x_2), \ldots (w_n, x_n)$ then $x_1, \ldots, x_n$ are marginally distributed according to $F_X$.

So to learn about $g(X)$, just throw away the samples of other variables.