Lec-42

$f: I \to \mathbb{R}$

Recall: ① Convex & Concave fn^{n}s, If f is a twice differentiable fn^n, then f is convex $\iff$ $f''(x) \geq 0 \ \forall x \in I$

f is concave $\iff$ $f''(x) \leq 0 \ \forall x \in I$

② Jensen's Inequality

③ Markov Inequality: $X \to$ non-neg. ran. var with finite n-th moment. Then for $\varepsilon > 0$

$$P\{X \geq \varepsilon\} \leq \frac{E(X^n)}{\varepsilon^n}$$

④ Chebyshev's Inequality: $X \to$ ran. var with finite mean μ and variance σ^2. Then for every $\varepsilon > 0$

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

Example ①: Let $X \sim B(n, p)$. Estimate $P(X \geq \alpha n)$, where $p < \alpha < 1$ using Markov (for first moment) and Chebyshev's inequality. Compare both the estimates for $p = \frac{1}{2}$ and $\alpha = \frac{3}{4}$.

Soln: Note that X takes values $\{0, 1, \ldots, n\}$, hence is a non-negative random variable and $E(X) = np$. Applying Markov's inequality, we obtain

$$P(X \geq \alpha n) \leq \frac{E(X)}{\alpha n} = \frac{pn}{\alpha n} = \frac{p}{\alpha}$$

Chebyshev's inequality gives estimate for $P\left(|X - E(X)| \geq \alpha n\right)$. So we have to rewrite the event $\{X \geq \alpha n\}$, so that we can use the Chebyshev's inequality.

$$P\{X \geq \alpha n\} = P\{X - np \geq \alpha n - np\}$$

$$\leq P\{|X - np| \geq \alpha n - np\} \quad \left[\because \{|Y| \geq \alpha\} = \{Y \leq -\alpha\} \cup \{Y \geq \alpha\}\right.$$

$|y| > \alpha$

$-\alpha \qquad \alpha$

$$\leq \frac{Var(X)}{(\alpha n - np)^2} = \frac{np(1-p)}{n^2(\alpha - p)^2}$$

$$= \frac{p(1-p)}{n(\alpha - p)^2}$$

By Markov's inequality for $p = \frac{1}{2}$ and $\alpha = \frac{3}{4}$, we have

$$P\left(X \geq \frac{3}{4}n\right) \leq \frac{2}{3}$$

By Chebyshev's inequality for $p = \frac{1}{2}$ and $\alpha = \frac{3}{4}$, we have

$$P\left(X \geq \frac{3n}{4}\right) \leq \frac{4}{n}$$

If $n \geq 6$, then estimate given by Chebyshev's are sharper than the estimates provided by Markov's.

Also as n increases, estimates given by Chebyshev's inequality decreases, i.e., gives much more information whereas the estimates provided by Markov inequality remains constant as n varies.

Law of Large Numbers

Theo: (Weak law of large numbers). Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean μ. Then for every $\delta > 0$,

$$\lim_{n\to\infty} P\left\{ \left| \frac{S_n}{n} - \mu \right| \geq \delta \right\} = 0 \quad \text{or equivalently}$$

$$\lim_{n\to\infty} P\left\{ \left| \frac{S_n}{n} - \mu \right| < \delta \right\} = 1 \qquad \text{—— (1)}$$

where $S_n = X_1 + X_2 + \cdots + X_n$

Remark: (1) The weak law of large numbers states that for large n, the bulk of the distribution of $\frac{S_n}{n}$ is concentrated near μ. That is, if we consider a positive length interval $[\mu-\delta, \mu+\delta]$ around μ, then there is a high probability that $\frac{S_n}{n}$ will fall in that interval; as $n\to\infty$, this probability converges to 1. Of course, if δ is very small we may have to wait longer (i.e, need a larger value of n) before we can assert that $\frac{S_n}{n}$ is highly likely to fall in that interval.

(2) To understand the convergence in weak law, think in terms of PMF (if X_i's are discrete random variables) or PDF (if X_i's have pdf then we know that S_n will possess a pdf as well) of random variable $\frac{S_n}{n}$.

Weak law states that "almost all" of the pmf or pdf of $\frac{S_n}{n}$ is concentrated within δ-neighborhood of μ for large values of n.

(3) The limit in Remark (1) means: $\forall\ \delta, \varepsilon > 0$, $\exists$ $n_0(\varepsilon, \delta)$ such that for all $n \geq n_0(\varepsilon, \delta)$, we have

$$P\left\{ \left| \frac{S_n}{n} - \mu \right| < \delta \right\} > 1 - \varepsilon.$$

If we refer to δ as the accuracy level and ε as the confidence level, the weak law takes the following intuitive form:

for any given level of accuracy and confidence, $\frac{S_n}{n}$ will be equal to μ, within these levels of accuracy and confidence, provided n is large enough.

Example (2): Let $X_1, X_2, \ldots$ be independent and identically distributed random variables with $E(X_i) = 0$ and $Var(X_i) = 1$ $\forall\ i$. Let $S_n = X_1 + X_2 + \cdots + X_n$.

Then for any $x > 0$, compute $\lim_{n \to \infty} P(-nx < S_n < nx)$.

Soln: For any $x > 0$, we have

$$P(-nx < S_n < nx) = P\left(-x < \frac{S_n}{n} < x\right) = P\left(\left| \frac{S_n}{n} - 0 \right| < x \right)$$

$$= 1 - P\left(\left|\frac{S_n}{n} - 0\right| \geq x\right).$$

By weak law of large numbers, we have

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - 0\right| \geq x\right) = 0$$

Random Sampling

Let $X_1, X_2, \ldots, X_n$ be n independent random variables having the same distribution. These random variables may be thought of as n independent measurements of some quantity that is distributed according to their common distribution. (E.g., height of students in LNMIIT). In this sense, we sometimes speak of the random variables $X_1, \ldots, X_n$ as constituting a random sample of size n from this distribution.

Suppose that the common distribution of these random variables have finite mean μ. Then for n sufficiently large, we would expect that the sample mean

$$\frac{S_n}{n} = \frac{X_1 + \cdots + X_n}{n}$$

should be close to the true mean μ.

The weak law of large numbers asserts that the sample mean of a large number of independent identically distributed random variables is very close to the true mean with high probability.

Weak law of large numbers says that for any $\delta > 0$,

$$\lim_{n \to \infty} P\left\{\left|\frac{S_n}{n} - \mu\right| \geq \delta\right\} = 0 \quad \text{—— (2)}$$

We may interpret eqn (2) in the following way:
The number δ can be thought of as the desired accuracy in the approximation of μ by $\frac{S_n}{n}$.

Eqn (2) assures us that no matter how small δ may be chosen, the probability that $\frac{S_n}{n}$ approximates μ to within this accuracy, that is, $P\left\{\left|\frac{S_n}{n}-\mu\right|<\delta\right\}$, converges to 1 as the no. of observations gets large.

<u>Example (4)</u>: <u>Probabilities and frequencies</u>: Consider an event A defined in the context of some probabilistic experiment. Let $p = P(A)$ be the probability of this event A. We consider n independent repetitions of this experiment and let M_n be the "fraction" of times that event A occurs in this context. M_n is often called the <u>empirical frequency</u> of A.

Note that $M_n = \dfrac{X_1 + \cdots + X_n}{n}$, where X_i is 1 whenever A occurs and 0 o.w.

In particular, $E(X_i) = p$.

The weak law of large numbers applies and shows that when n is large, the empirical frequency is most likely to be within ε of p.

Loosely speaking, this allows us to conclude that empirical frequencies are faithful estimates of p. Alternatively, this is a step towards interpreting the probability p as the frequency of the occurrence of A.