Limit Theorems [Ref. [Sheldon M. Ross]]

March 26, 2020
Dr. SUSHIL KUMAR AZAD

The limit theorems, such as laws of large numbers or Central limit theorems are the most important results in probability theory. We shall start this lecture by proving a result known as Markov's inequality.

Theorem 1. (Markov's inequality) If X is a random variable that takes only nonnegative values, then for any value $a>0$,

$$\boxed{P\{X \geq a\} \leq \frac{E[X]}{a}}$$

Proof. For $a>0$, let us define a random variable, Y, by

$$Y = \begin{cases} 1 & \text{if } X \geq a \\ 0 & \text{otherwise} \end{cases}$$

Since $X \geq 0$, we have $Y \leq \frac{X}{a}$

$$\Rightarrow E[Y] \leq \frac{E[X]}{a}$$

But, $E[Y] = P\{X \geq a\}$

$$\Rightarrow P\{X \geq a\} \leq \frac{E[X]}{a},$$ which is the result. //

SUSHIL KUMAR AZAD

Alternative proof of Markov's inequality, assuming, X is a continuous r.v. with pdf $f(x)$, $x \geq 0$. Then

$$E[X] = \int_0^\infty x f(x)\,dx = \int_0^a x f(x)\,dx + \int_a^\infty x f(x)\,dx$$

$$\Rightarrow E[X] \geq \int_a^\infty x f(x)\,dx, \quad \left(\text{as } \int_0^a x f(x)\,dx \geq 0\right)$$

$$\geq \int_a^\infty a f(x)\,dx = a\int_a^\infty f(x)\,dx$$

$$\Rightarrow E(X) \geq a\,P\{X \geq a\} \text{ or}$$

$$P\{X \geq a\} \leq \frac{E[X]}{a}.$$

Remark. The Markov's inequality enable us to derive bounds on probabilities when only the mean of the probability distribution is known.

Next, we obtain Chebyshev's inequality, as a corollary to Theorem 1, which enable us to derive bounds on probabilities when both the mean and the variance of the probability distribution are known.

SUSHIL KUMAR AZAD

Theorem 2. (Chebyshev's Inequality) If X is a random variable with finite mean, μ and variance, σ^2, then for any value $k > 0$

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

Proof. Let $Y = (X-\mu)^2$, then Y is a nonnegative random variable. Let $a = k^2 > 0$.

Using Markov's inequality we get

$$P\{Y \geq a\} \leq \frac{E[Y]}{a} \quad \text{----} (i)$$

Now $Y \geq a$, i.e., $(X-\mu)^2 \geq k^2 \Leftrightarrow |X-\mu| \geq k$.

$$\therefore (i) \Rightarrow P\{|X-\mu| \geq k\} \leq \frac{E[(X-\mu)^2]}{a}$$

$$= \frac{\sigma^2}{k^2}, \quad \left(\sigma^2 \text{ variance } = E[(X-\mu)^2], \text{ by definition}\right)$$

i.e. $P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$, which is the result. //

Remark. The upper bound that the Chebyscher's inequality provides is not necessarily be very close to the actual probability in most cases.

Example 28. Let $X \sim N(\mu, \sigma^2)$. Find the upper bound of probability that $|X-\mu|$ exceeds the twice of the standard deviation of X, using Chebyscher's inequality. Also, Compare it with the actual probability.

Solution. Using the Chebyscher's inequality, we get

$$P\{|X-\mu| > 2\sigma\}$$

$$\leq \frac{\sigma^2}{4\sigma^2} \quad (\text{Here, } k = 2\sigma)$$

$$= \frac{1}{4} = 0.25$$

i.e. $P\{|X-\mu| > 2\sigma\} \leq 0.25$

On the other hand, using the definition of standard normal distribution, we get

$$P\{|X-\mu| > 2\sigma\}$$

$$= P\left\{\left|\frac{X-\mu}{\sigma}\right| > 2\right\}$$

$$= P\{|Z| > 2\} \quad \left(Z = \frac{X-\mu}{\sigma} \text{ is a standard normal variable}\right)$$

$$= 2[1 - \Phi(2)]$$

$$= 0.0456$$

Note that the actual probability, 0.0456 is not close to the upper bound, 0.25, obtained of Chebyschv's inequality. //

Example 29. Let X be uniformly distributed over $(0, 10)$. Obtain

(a) the lower bound for the $P\{1 \le X \le 9\}$, using Chebyschev's inequality

(b) Compare the answer of part (a) with actual probability.

Solution. (a) We know that if $X \sim U(a, b)$, then

$$E(X) = \frac{b+a}{2}, \text{ and } var(X) = \frac{(b-a)^2}{12}$$

So, we have $E(X) = \frac{10+0}{2} = 5$, and

$$var(X) = \frac{(10-0)^2}{12} = \frac{100}{12} = \frac{25}{3}$$

i.e, we have $\mu = 5$ and $\sigma^2 = \frac{25}{3}$.

Now $P\{1 \le X \le 9\}$

$$= P\{-4 \le X-5 \le 4\}$$

$$= P\{|X-5| \le 4\}$$

$$= 1 - P\{|X-5| > 4\}$$

$$\ge 1 - \frac{\sigma^2}{(16)}, \text{ using Chebyschev's inequality}$$

$$= 1 - \frac{25}{3 \times 16}$$

$$\simeq 0.52$$

i.e., $P\{1 \le X \le 9\} \ge 0.52$

$\Rightarrow$ the lower bound of the given probability is 0.52.

SUSHIL KUMAR AZAD

(b) $P\{1 \le X \le 9\} = \int_1^9 f(x)\,dx$,

where $f(x) = \begin{cases} \frac{1}{10} & \text{if } 0 < x < 10 \\ 0 & \text{otherwise.} \end{cases}$

is the pdf of a uniform random variable, X, on the interval $(0, 10)$.

$\therefore P\{1 \le X \le 9\} = \int_1^9 \frac{1}{10}\,dx$

$= \frac{8}{10} = 0.80$

Clearly $0.80 > 0.52$, that is, the actual probability is not close to the lower bound obtained by Chebyschev's inequality.

Remark. If X is a random variable with finite mean, μ and variance, σ^2, then for any value $K > 0$,

$$\boxed{P\{|X - \mu| < K\} \ge 1 - \frac{\sigma^2}{K^2}},$$

which follows from Chebyschev's inequality.

Example 30. Let $X_1, X_2, \ldots, X_{10}$ be independent Poisson random variables with mean 1. Find a bound on

$P\{X_1 + \cdots + X_{10} \ge 15\}$, using

the Markov's inequality.

SUSHIL KUMAR AZAD

SUSHIL KUMAR AZAD

Solution (Example 30)

(a) We have $X_i \in P(\lambda)$ $\forall i = 1, 10$ where λ, mean $= 1$.
Also, since the variance of the Poisson variate is $= \lambda \Rightarrow var(X_i) = 1$ $\forall i = 1, 10$

So, $E\left[\frac{X_1 + X_2 + \cdots + X_{10}}{10}\right] = 1$, and

$$var\left[\frac{X_1 + X_2 + \cdots + X_{10}}{10}\right] = \frac{1}{10}$$

Hence, by Markov's inequality, we get

$$P\{|X_1 + \cdots + X_{10}| \geq 15\} = P\{X_1 + \cdots + X_{10} \geq 15\}$$ as $X_i \in P(1) \geq 0$

$$= P\left\{\left(\frac{X_1 + \cdots + X_{10}}{10}\right) \geq 1.5\right\}$$

$$\leq \frac{E\left[\frac{X_1 + \cdots + X_{10}}{10}\right]}{1.5} = \frac{1}{1.5} = \frac{2}{3}$$

i.e. $P\{X_1 + \cdots + X_{10} \geq 15\} \leq \frac{2}{3}$.

Now, we shall prove results which are based on Chebyschev's inequality.

Theorem 3. If $var(X) = 0$, then

$$P\{X = E(X)\} = 1,$$

that is, the only random variables having variances equal to 0 are those that are constant with probability 1.

Proof. Let $k = \frac{1}{n}$, where $n \geq 1$. Then, by Chebyshev's inequality

$$P\{|X - \mu| > \frac{1}{n}\} \leq n^2\sigma^2 = 0, \text{ as } \sigma^2 = 0$$

$$\Rightarrow P\{|X - \mu| > \frac{1}{n}\} = 0, \text{ for any } n \geq 1.$$

Continued..

Thus, $0 = \lim_{n \to \infty} P\{|X-\mu| > \frac{1}{n}\} = P\{\lim_{n \to \infty}\{|X-\mu| > \frac{1}{n}\}\}$

$= P\{|X-\mu| > 0\}$

i.e. $0 = P\{X \neq \mu\}$

$\Rightarrow P\{X = \mu\} = 1 - P\{X \neq \mu\} = 1$, which is the result. //

Theorem 3. (The 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, and finite variance, σ^2. Then, for any $k > 0$,

$$P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq k\right\} \to 0 \text{ as } n \to \infty$$

SUSHIL KUMAR AZAD

Proof. We have

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n}\left[E(X_1) + \cdots + E(X_n)\right] = \frac{n\mu}{n} = \mu, \text{ and}$$

$$\operatorname{var}\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{var}(X_i) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Hence, from Chebyschev's inequality, we get

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq k\right\} \leq \frac{\sigma^2}{nk^2},$$

where the R.H.S. $\to 0$ as $n \to \infty$,

and so

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq k\right\} \to 0 \text{ as } n \to \infty$$

which is the result //

Remark. The weak law of large numbers states that for any specified large value, m,

$$\left(\frac{X_1 + \cdots + X_m}{m}\right)$$ is likely to be near μ,

However, it does not say that

$$\left(\frac{X_1 + \cdots + X_n}{n}\right)$$ is bound to stay near μ

for all values of $n \geq m$.

In other words, it leaves open the possibility that large values of

$$\left|\left(\frac{X_1 + \cdots + X_n}{n}\right) - \mu\right|$$

can occur infinitely often, though at infrequent intervals. //.

Now, we shall state the best-known result in probability theory, viz., the Strong law of large numbers, which shows that the large values of $\left|\left(\frac{X_1 + \cdots X_n}{n}\right) - \mu\right|$ cannot occur. In fact, it states that with probability 1, for any positive value ε,

$$\left|\sum_{i=1}^{n} \frac{X_i}{n} - \mu\right|$$

will be greater than ε only a finite number of times.

Theorem 4 Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having a finite mean μ $(= E(X_i))$. Then,

continued.

SUSHIL KUMAR AZAD

SUSHIL KUMAR AZAD

with probability 1,

$$\frac{\sum_{i=1}^{n} X_i}{n} \longrightarrow \mu \quad \text{as } n \to \infty$$

Remark. The above theorem 4, was originally proved, in the special case of Bernoulli random variables, by the French mathematician, Borel. The general form of this was proved by the Russian mathematician A.N. Kolmogorov.

Example 31 (application of Theorem 4)

Suppose that a sequence of independent trials is performed. Let E be a fixed event, and $P(E)$ is the probability that E occurs on any particular trial. Then, with probability 1, the limiting proportion of time that the event E occurs is just $P(E)$.

Solution. Let us define the r.v.'s

$$X_i = \begin{cases} 1 & \text{if } E \text{ occurs on the } i\text{th trial} \\ 0 & \text{if } E \text{ does not occur on the } i\text{th trial} \end{cases}$$

Then, by the strong law of large numbers, with probability 1, we have

$$\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mu = E(X_i) \quad \text{as } n \to \infty$$

But, $E(X_i) = P(E) \ \forall i$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} X_i \to P(E) \quad \text{as } n \to \infty \quad \text{(ii)}$$

Now, $(X_1 + \cdots + X_n)$ represents the number of times that the event E occurs in the first n trials, hence (ii) $\Rightarrow$ with probability 1, the limiting proportion of time that the event E occurs is just $P(E)$. //