4.2

Convergence of Sequence of Random Variables

In this module we investigate convergence properties of sequences of random variables. Throughout this module we assume that $\{X_1, X_2, \dots\}$ or $\{X_n\}$ is a sequence of r.vs and X is a r.v. We consider ***four different modes of convergence for random variables.***

1. **Almost sure convergence**: It is the **probabilistic version of pointwise convergence** known from elementary real analysis. It is also known as **convergence with probability one.**

   The sequence of r.vs $\{X_n\}$ is said to ***converge almost surely*** to a r.v. X if

   $$P\left(\left\{w : \lim_{n\to\infty} X_n(w) = X(w)\right\}\right) = 1$$

   In this case we write $\boldsymbol{X_n \xrightarrow{a.s} X}$ (or $\boldsymbol{X_n \longrightarrow X}$ **with probability** $\mathbf{1}$).

2. **Convergence in probability:** It is essentially mean that the probability that $|X_n - X|$ exceeds any prescribed strictly positive value, converges to zero. The basic idea behind this type of convergence is that the probability of an *unsual* outcome becomes smaller and smaller as the sequence progresses. The sequence of r.vs $\{X_n\}$ is said to ***converge in probability*** to a r.v. X if

   $$\lim_{n\to\infty} P(\{|X_n - X| > \epsilon\}) = 0$$

   for every $\epsilon > 0$.It is denoted by $X_n \xrightarrow{\boldsymbol{P}} X$.

3. **Convergence in $\boldsymbol{r}^{\text{th}}$ mean:** Let $\{X_n\}$ be a sequence of r.vs such that $E(|X_n|^r) < \infty$ for some $r > 0$. We say that X_n ***converges in the $\boldsymbol{r^{th}}$ mean*** to a r.v. X if $E(|X|^r) < \infty$ and

   $$E(|X_n - X|^r) \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

   and we write $\boldsymbol{X_n \xrightarrow{r} X.}$

If $r = 2$, we call it as **convergence in quadratic mean** and it is denoted by $\boldsymbol{X_n \xrightarrow{q.m} X}$

4. **Convergence in distribution: Convergence in distribution** is very frequently used in practice, most often it arises from the application of the **central limit theorem** (to be discussed in module 4.5). Note that a cumulative distribution function (c.d.f) is briefly called as *distribution function (d.f)* also.

Let $\{F_n\}$ be a sequence of cumulative distribution functions (c.d.fs), if there exists a c.d.f. F such that as $n \to \infty$,

$$F_n(x) \longrightarrow F(x)$$

for all x at which F is continuous, then we say that F_n **converges weakly** to F, and it is denoted by $\boldsymbol{F_n \xrightarrow{w} F}$.

If $\{X_n\}$ is a sequence of r.vs and $\{F_n\}$ is the corresponding sequence of c.d.fs, then we say that $\boldsymbol{X_n}$ ***converges in distribution*** (or **law**) to $\boldsymbol{X}$ if there exists an r.v X with c.d.f. F such that $F_n \xrightarrow{w} F$. We write $X_n \xrightarrow{d} X$ or $X_n \xrightarrow{L} X$.

Note: It is quite possible for a given sequence of c.d.fs to converge to a function that is not a c.d.f.

Example: Let $F_n(x) = \begin{cases} 0\ , x < n \\ 1\ , x \geq n \end{cases}$

As $n \to \infty$, $F_n(x) \longrightarrow F(x) = 0$ which is not a c.d.f.

Example 1: **Let $\boldsymbol{X_1, X_2, \ldots, X_n}$ be i.i.d.r.vs with common p.d.f**

$$\boldsymbol{f(x)} = \begin{cases} \boldsymbol{\frac{1}{\theta}} & , \ \boldsymbol{0} < x < \theta\ , \theta > 0 \\ \boldsymbol{0} & , \ \boldsymbol{otherwise} \end{cases}$$

Let $\boldsymbol{X_{(n)} = max(X_1, \ldots, X_n)}$. Then show that $\boldsymbol{X_{(n)} \xrightarrow{L} X}$, where $\boldsymbol{X}$ is degenerate at $\boldsymbol{x = \theta}$.

(Note: We say that a r.v.X is **degenerate at $\boldsymbol{x = \theta}$** if $P(X = \theta) = 1$)

Solution: Corresponding to p.d.f. $f(x) = \frac{1}{\theta}$, the c.d.f. is given by

$$F(x) = \int_0^x f(t)dt = \frac{1}{\theta}\int_0^x dt = \frac{x}{\theta}$$

$$\Rightarrow F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x}{\theta} & , \quad 0 \le x < \theta \\ 1 & , \quad x \ge \theta \end{cases}$$

Then the c.d.f. of $X_{(n)}$ is given by

$$F_n(x) = [F(x)]^n = \begin{cases} 0 & , \quad x < 0 \\ \left(\frac{x}{\theta}\right)^n & , \quad 0 \le x < \theta \\ 1 & , \quad x \ge \theta \end{cases}$$

We see that as $n \longrightarrow \infty$

$$F_n(x) = F(x) = \begin{cases} 0 \ \ if \ \ x < \theta \\ 1 \ \ if \ \ x \ge \theta \end{cases}$$

which is the d.f. of $P(X = \theta) = 1.\ i.e.$, X is degenerate at $x = \theta$.

Thus $F_n \xrightarrow{w} F$and hence $X_n \xrightarrow{L} X$.

The following example shows that convergence in distribution does not imply convergence of moments.

Example 2: Let F_n be a sequence of c.d.fs defined by

$$\boldsymbol{F_n(x)} = \begin{cases} \mathbf{0} & , \quad \boldsymbol{x} < 0 \\ \mathbf{1} - \frac{\mathbf{1}}{\boldsymbol{n}} & , \quad \mathbf{0} \le \boldsymbol{x} < n \\ \mathbf{1} & , \quad \boldsymbol{x} \ge \boldsymbol{n} \end{cases}$$

Show that $X_n \xrightarrow{L} X$ does not imply $E\left({X_n}^k\right) \longrightarrow E\left(X^k\right)$.

Solution: We see that as $n \longrightarrow \infty$

$$F(x) = \begin{cases} 0 \ , \ x < 0 \\ 1 \ , \ x \geq 0 \end{cases}$$

Note that F_n is the c.d.f. of the r.v. X_n with p.m.f.

$$P(X_n = 0) = 1 - \frac{1}{n} \ , P(X_n = n) = \frac{1}{n}$$

and F is the c.d.f. of the r.v. degenerate at 0 $i.e.$, $P(X = 0) = 1$.

Thus, $F_n \xrightarrow{w} F$ and hence $X_n \xrightarrow{L} X$. We have

$E\left({X_n}^k\right) = 0^k \left(1 - \frac{1}{n}\right) + n^k \left(\frac{1}{n}\right) = n^{k-1}$, where k is a positive integer. Also, $E(X^k) = 0^k 1 = 0$. Hence $E\left({X_n}^k\right) \not\rightarrow E(X^k)$ as $n \longrightarrow \infty$

Therefore, $X_n \xrightarrow{L} X$does not imply $E\left({X_n}^k\right) \longrightarrow E(X^k)$.

The next example shows that weak convergence of distribution of function does not imply the convergence of corresponding p.m.fs or p.d.fs.

Example 3: Let $\{X_n\}$ be a sequence of r.vs with p.m.f.

$$\boldsymbol{f_n(x) = P(X_n = x) = \begin{cases} 1 \ , \ if \quad x = 2 + \frac{1}{n} \\ 0 \ , \quad otherwise \end{cases}}$$

Show that $F_n \xrightarrow{w} F$ does not imply $\boldsymbol{f_n \longrightarrow f}$.

Solution: Note that $f_n(x) \longrightarrow f(x)$ as $n \longrightarrow \infty$, where $f(x) = 0$ for all x.

The c.d.f. of X_n is given by

$$F_n(x) = P(X_n \leq x) = \begin{cases} 0 \ , \ \ x < 2 + \frac{1}{n} \\ 1 \ , \ \ x \geq 2 + \frac{1}{n} \end{cases}$$

which converges to

$$F(x) = \begin{cases} 0 & , \quad x < 2 \\ 1 & , \quad x \geq 2 \end{cases}$$

at all continuity points of F. Since F is the c.d.f. of a r.v. degenerate at $x = 2$ $i.e.$, $P(X = 2) = 1$

$$i.e\ , f(x) = \begin{cases} 1\ , & x = 2 \\ 0\ , & otherwise \end{cases}$$

Thus, convergence of distribution functions does not imply the convergence of corresponding p.m.fs.

Example 4: Let $\{X_n\}$ be a sequence of r.vs with p.m.f $P(X_n = 1) = \frac{1}{n}$ and $P(X_n = 0) = 1 - \frac{1}{n}$. Then show that $X_n \xrightarrow{P} 0$.

Solution: We have $P(|X_n| > \epsilon) = \begin{cases} P(X_n = 1) = \frac{1}{n}\ , & 0 < \epsilon < 1 \\ 0 & , \quad \epsilon \geq 1 \end{cases}$

It follows that $P(|X_n| > \epsilon) \longrightarrow 0$ as $n \longrightarrow \infty$, and we conclude that $X_n \xrightarrow{P} 0$

Example 5: Let $\{X_n\}$ be a sequence of r.vs defined by

$$\boldsymbol{P(X_n = 0) = 1 - \frac{1}{n},\ P(X_n = 1) = \frac{1}{n}\ , n = 1, 2, \ldots}$$

Show that $X_n \xrightarrow{q.m} X$, where $P(X = 0) = 1$.

Solution: Consider $E(|X_n - 0|^2) = E(|X_n|^2) = E\left({X_n}^2\right) = 0^2\left(1 - \frac{1}{n}\right) + 1^2\left(\frac{1}{n}\right)$

$$= \frac{1}{n} \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

Thus, $X_n \xrightarrow{q.m} X$, where X is degenerate at 0.

Example 6: Let $\{X_n\}$ be a sequence of independent r.vs defined by

$$P(X_n = 0) = 1 - \frac{1}{n} \text{ and } P(X_n = 1) = \frac{1}{n}, \quad n = 1, 2, \ldots$$

Show that $X_n \xrightarrow{q.m} 0$ but $X_n \stackrel{a.s}{\not\longrightarrow} 0$

Solution: $E(|X_n - 0|^2) = E(|X_n|^2) = 0^2\left(1 - \frac{1}{n}\right) + 1^2\left(\frac{1}{n}\right) = \frac{1}{n} \longrightarrow 0$ as $n \longrightarrow \infty$

Hence $X_n \xrightarrow{q.m} 0$.

Also, $P(X_n = 0 \ \ for\ every\ m \le n \le n_0) = \prod_{n=m}^{n_0}\left(1 - \frac{1}{n}\right) = \frac{m-1}{n_0}$ which converges to zero as $n \longrightarrow \infty$ for all values of m. Thus, $X_n \stackrel{a.s}{\not\longrightarrow} 0$

Example 7: Let $\{X_n\}$ be a sequence of independent r.vs defined by

$$P(X_n = 0) = 1 - \frac{1}{n^r} \text{ and } P(X_n = n) = \frac{1}{n^r}, r \ge 2, \ n = 1, 2, \ldots$$

Show that $X_n \xrightarrow{a.s} 0$ but $X_n \stackrel{r}{\not\longrightarrow} 0$.

Solution: We have $P(X_n = 0 \ \ for \ \ m \le n \le n_0) = \prod_{n=m}^{n_0}\left(1 - \frac{1}{n^r}\right)$

As $n_0 \longrightarrow \infty$, the infinite product converges to some nonzero quantity, which itself converges to 1 as $m \longrightarrow \infty$.

That is, $P\left[\lim_{n\to\infty} X_n = 0\right] = 1$. Therefore $X_n \xrightarrow{a.s} 0$

However, $E(|X - 0|^r) = E(|X|^r) = 0^r\left(1 - \frac{1}{n^r}\right) + n^r \times \frac{1}{n^r} = 1$

and hence $E(|X|^r) = 1$ as $n \longrightarrow \infty$. Therefore, $X_n \stackrel{r}{\not\longrightarrow} 0$

Thus, $X_n \xrightarrow{a.s} 0$ but $X_n \stackrel{r}{\not\longrightarrow} 0$

A sufficient condition for $a.s.$ convergence:

We state a sufficient condition for the $a.s.$ convergence without proof which is sometimes to verify.

$$X_n \xrightarrow{a.s} X \Leftrightarrow \lim_{n\to\infty} P\left[\bigcup_{m=n}^{\infty} |X_n - X| > \epsilon\right] = 0, \qquad \forall\, \epsilon > 0$$

Example 8: Let $\{X_n\}$ be a sequence of r.vs with $P\left(X_n = \pm\frac{1}{n}\right) = \frac{1}{2}$. Show that $X_n \xrightarrow{r} 0$ and $X_n \xrightarrow{a.s} 0$.

Solution: We have $E(|X_n - 0|^r) = E(|X_n|^r) = \frac{1}{n^r}\left(\frac{1}{2}\right) + \frac{1}{n^r}\left(\frac{1}{2}\right) = \frac{1}{n^r} \longrightarrow 0$ as $n \longrightarrow \infty$ and hence $X_n \xrightarrow{r} 0$. It follows that

$$\bigcup_{j=n}^{\infty}\left\{\left|X_j\right| > \varepsilon\right\} = \left\{\left|X_n\right| > \varepsilon\right\}$$

Choosing $n > \frac{1}{\epsilon}$, we see that

$$P\left[\bigcup_{j=n}^{\infty}\left\{\left|X_j\right| > \varepsilon\right\}\right] = P\left(\left\{\left|X_n\right| > \varepsilon\right\}\right) \le P\left(|X_n| > \frac{1}{n}\right) = 0 \text{ as } n \longrightarrow \infty$$

$$\Rightarrow \lim_{n\to\infty} \mathrm{P}\left[\bigcup_{j=n}^{\infty}\left\{\left|\mathrm{X}_j\right| > \varepsilon\right\}\right] = 0 \Longrightarrow X_n \xrightarrow{a.s} 0$$

Implications always valid between modes of convergence

We state the following implications always valid between modes of convergence without proof.

1) $X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$
2) $X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

Counter examples to implications among the modes of convergence

1) $X_n \xrightarrow{d} X \not\Rightarrow X_n \xrightarrow{P} X$ (**See P1**)
2) $X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{r} X$ (**See P2**)
3) $X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{a.s} X$ (**See P3**)
4) $X_n \xrightarrow{r} X \not\Rightarrow X_n \xrightarrow{a.s} X$
5) $X_n \xrightarrow{a.s} X \not\Rightarrow X_n \xrightarrow{r} X$

The following theorem is known as **Slutsky's Theorem** and is very useful in finding the limiting distribution of certain r.vs. This theorem is stated without proof.

Theorem 1: Slutsky's Theorem: Let $\{X_n, Y_n\}, n = 1, 2, \ldots$ be a sequence of pairs of random variables and let c be a constant. If $X_n \xrightarrow{L} X$ and $Y_n \xrightarrow{P} c$, then

(i) $X_n + Y_n \xrightarrow{L} X + c$

(ii) $X_n Y_n \xrightarrow{L} cX$

(iii) $\frac{X_n}{Y_n} \xrightarrow{L} \frac{X}{C}$ **if** $c \neq 0$

An example presented in **P4** as an application of **Slutsky's theorem.**