UDC: 517.521
Review article

THE RELATIONS BETWEEN MODES OF CONVERGENCE FOR SEQUENCES OF RANDOM VARIABLES

Bedrije BEXHETI[1*], Alit IBRAIMI[1], Flamure SADIKI[1], Ferzije LLESHI POLLOZHANI[1]

[1] Department of Natural Sciences and Mathematics, University of Tetovo, NMK
[*]Corresponding author email: bedrije.bexheti@unite.edu.mk

Abstract

In this paper, we are going to analyze the relations between different types of convergence of a random sequence, such as almost sure convergence, convergence in mean square, convergence in distribution and convergence in probability. The convergence in distributions says nothing about the relationship between the random variables X_n and X, while for convergence in probability, the joint distribution of X_n and X is relevant.
In the main part of the paper, we are going to prove the theorem which argues that the convergence in probability implies convergence in distribution, and the opposite is not true. But if $X_n \xrightarrow{d} c$, where c is a constant, then $X_n \xrightarrow{p} c$, which mean that convergence in probability to a constant is equivalent to convergence in distributions.
Also, we give some interesting examples.

Keywords: random variable, random sequence, mean square, convergence in distributions.

1. Introduction

One of the most important parts of probability theory concerns the behavior of sequences of random variables. This part of probability is often called “limit theory” or “asymptotic theory.” This concept is very important for statistical inference. Since statistics is all about gathering data, we will naturally be interested in what happens as we gather more and more data, hence our interest in this question. The paper develops appropriate methods of discussing convergence of random variables.

In the first part we will give definitions of different types of convergence and discuss how they are related. A sequence might converge in one sense but not another. Some of these convergence types are "stronger" than others and some are "weaker." Consider a sequence of random variables $X_1, X_2, \ldots$. This sequence might "converge" to a random variable X . There are four types of convergence that we will discuss in this part:

- Almost sure convergence.
- Convergence in mean
- Convergence in distribution,
- Convergence in probability.

An infinite sequence $X_n, n = 1, 2, \ldots$, of random variable is called a random sequence.

2. Different Types of Convergence for Sequences of Random Variables

2.1. Almost Sure Convergence:

Definition 2.1.1. A sequence of random variables $X_1, X_2, \ldots$ converges **almost surely,** or with probability one, to the random variable X, shown by $X_n \xrightarrow{a.s} X$, if

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

In some problems, proving almost sure convergence directly can be difficult. Thus, it is desirable to know some sufficient conditions for almost sure convergence. Here is a result that is sometimes useful when we would like to prove almost sure convergence.

Theorem 2.1.1. Consider the sequence $X_1, X_2, \ldots$. For any $\varepsilon > 0$, define the set of events

$$A_m = \left\{\left|X_n - X\right| < \varepsilon, \text{ for all } n \geq m\right\}.$$

Then $X_n \xrightarrow{a.s} X$ if and only if for any $\varepsilon > 0$, we have

$$\lim_{m\to\infty} P(A_m) = 1.$$

An important example for almost sure convergence is the strong law of large numbers.

Theorem 2.1.3. *(Strong law of large numbers)*
Let $X_1, X_2, \ldots$ be i.i.d random variable with a finite expected value $E\left[X_i\right] = \mu < \infty, \forall i.$ Let

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Then $M_n \xrightarrow{a.s} \mu.$

2.2. Convergence in Mean:

Definition 2.2.1. Let $r \geq 1$ be a fixed number. A sequence of random variables $X_1, X_2, \ldots$ converges in the ***r*-th mean** or in the L^r norm to a random variable X, shown by $X_n \xrightarrow{L^r} X$, if

$$\lim_{n\to\infty} E\left[\left|X_n - X\right|^r\right] = 0.$$

If $r = 2$, it is called the **mean-square convergence**, and it is shown by $X_n \xrightarrow{m.s} X$.

Theorem 2.2.1. Let $1 \le r \le s$. If $X_n \xrightarrow{L^s} X$, then $X_n \xrightarrow{L^r} X$.

2.3. Convergence in Distribution:

Convergence in distribution is in some sense the weakest type of convergence.

Definition 2.3.1 A sequence of random variables $X_1, X_2, \ldots$ converges **in distribution** to a random variable X, shown by $X_n \xrightarrow{d} X$, if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x),$$

for all x at which $F_X(x)$ is continuous.

2.4. Convergence in probability:

Definition 2.4.1. A sequence of random variables $X_1, X_2, \ldots$ converges **in probability** to a random variable X, shown by $X_n \xrightarrow{p} X$, if

$$\lim_{n\to\infty} P\left(\left|X_n - X\right| \ge \varepsilon\right) = 0, \text{ for all } \varepsilon > 0.$$

The most famous example of convergence in probability is the weak law of large numbers (WLLN)

Theorem 2.4.1. *(Weak law of large numbers).* If $X_1, X_2, \ldots$ are independent and identically distributed random variables with mean $E\left[X_i\right] = \mu < \infty, \forall i$, then the average sequence defined by

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

converges in probability to μ. It is called the "weak" law because it refers to convergence in probability.

3. Main Results

If a sequence of random variables converges *in probability* to a random variable X, then the sequence converges *in distribution* to X as well.

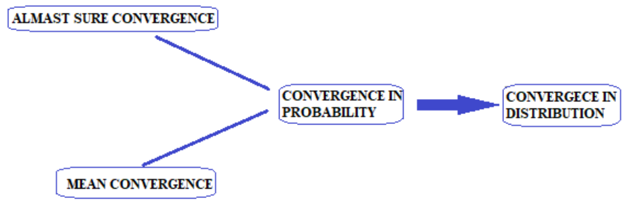

Fig 2.1. Relations between different types of convergence

Convergence in mean and almost sure convergence are stronger than convergence in probability.
There is no relation between almost sure and mean square convergence.

Theorem 3.1. The following relationships hold:

a) $X_n \xrightarrow{a.s} X$ implies that $X_n \xrightarrow{p} X$.

b) $X_n \xrightarrow{L^r} X$ implies that $X_n \xrightarrow{p} X$.

c) $X_n \xrightarrow{p} X$ implies that $X_n \xrightarrow{d} X$.

Proof. a) For any $\varepsilon > 0$,

$$\left[X_n \text{ converges to } X\right] \subseteq \bigcup_{n=1}^{\infty} B_n(\varepsilon)$$

With monotone increasing events

$$B_n(\varepsilon) = \bigcap_{m=n}^{\infty}\left[\left|X_m - X\right| \le \varepsilon\right], n = 1, 2, \ldots$$

Therefore,

$$P\left[X_n \text{ converges to X}\right] \le \lim_{n\to\infty} P\left[B_n(\varepsilon)\right]$$

By monotonicity!

If $P\left[X_n \text{ converges to X}\right] = 1$, then $\lim_{n\to\infty} P\left[B_n(\varepsilon)\right] = 1$ becomes

$$0 = \lim_{n\to\infty} P\left[B_n(\varepsilon)^c\right] \ge \lim_{n\to\infty} P\left[\bigcup_{m=n}^{\infty}\left[\left|X_m - X\right| > \varepsilon\right]\right]$$

By complementarity, whence

$$\lim_{n\to\infty} P\left[\left|X_n - X\right| > \varepsilon\right] = 0$$

Another proof;

The alternative definition of convergence in probability is given by

$$P\left(\left|X_n - X\right| \le \varepsilon\right) \to \infty \text{ as } n \to \infty$$

The alternative definition of almost convergence is given by

$$\forall \varepsilon > 0, P\left(\left|X_k - X\right| \le \varepsilon \text{ for all } k \ge n\right) \to 1 \text{ as } n \to \infty$$

Then

$$\begin{aligned}
&\forall \varepsilon > 0, \\
&1 \ge P\left(\left|X_n - X\right| \le \varepsilon\right) \ge P\left(\left|X_k - X\right| \le \varepsilon \text{ for all } k \ge n\right) \\
&\qquad\qquad \to 1 \text{ as } n \to \infty \text{ (because } X_n \xrightarrow{a.s} X) \\
&\text{So } X_n \xrightarrow{p} X
\end{aligned}$$

b) For any $\varepsilon > 0$, we have

$$\begin{aligned}
P\left(\left|X_n - X\right| \ge \varepsilon\right) &= P\left(\left|X_n - X\right|^r \ge \varepsilon^r\right) \text{(since } r \ge 1) \\
&\le \frac{E\left|X_n - X\right|^r}{\varepsilon^r} \text{(by Markov's inequality).}
\end{aligned}$$

Since by assumption $\lim_{n\to\infty} E\left(\left|X_n - X\right|^r\right) = 0,$ we conclude

$$\lim_{n\to\infty} P\left(|X_n - X| \geq \varepsilon\right) = 0, \text{for all } \varepsilon > 0.$$

c) This proof is a little more complicated.
Fix $\varepsilon > 0$. Then

$$F_n(x) = P(X_n \leq x) = P\left(X_n \leq x, X_n \leq x+\varepsilon\right) + P\left(X_n \leq x, X_n > x+\varepsilon\right)$$
$$\leq P\left(X \leq x+\varepsilon\right) + P\left(|X_n - x| > \varepsilon\right)$$
$$= F(x+\varepsilon) + P\left(|X_n - x| > \varepsilon\right).$$

Also,

$$F(x-\varepsilon) = P\left(X \leq x-\varepsilon\right) = P\left(X \leq x-\varepsilon, X_n \leq x\right) + P\left(X \leq x+\varepsilon, X > x\right)$$
$$\leq F_n(x) + P\left(|X_n - x| > \varepsilon\right).$$

Hence,

$$F(x-\varepsilon) - P\left(|X_n - X| > \varepsilon\right) \leq F_n(x) \leq F(x+\varepsilon) + P\left(|X_n - X| > \varepsilon\right).$$

Take the limit as $n \to \infty$ to conclude that

$$F(x-\varepsilon) \leq \liminf_{n\to\infty} F_n(x) \leq \limsup_{n\to\infty} F_n(x) \leq F(x+\varepsilon).$$

This holds for all $\varepsilon > 0$. Take the limit as $\varepsilon \to 0$ and use the fact that F is continuous at *x* and conclude that $\lim_{n\to\infty} F_n(x) = F(x)$.

The converse of Theorem 3.1. a), b) is not true in general. That is, there are sequences that converge in probability but not in mean or almost sure.

Convergence in probability is stronger than convergence in distribution. That is, if $X_n \xrightarrow{p} X$, then $X_n \xrightarrow{d} X$. The converse is not necessarily true.

A special case in which the converse is true is when $X_n \xrightarrow{d} c$, where *c* is a constant. In this case, convergence in distribution implies convergence in probability. We can state the following theorem:

Theorem 3.2. If $X_n \xrightarrow{d} c$, where c is a constant, then $X_n \xrightarrow{p} c$

Proof. Since $X_n \xrightarrow{d} c$, we conclude that for any $\varepsilon > 0,$ we have

$$\lim_{n\to\infty} F_{X_n}\left(c-\varepsilon\right) = 0,$$

$$\lim_{n\to\infty} F_{X_n}\left(c+\frac{\varepsilon}{2}\right) = 1.$$

We can write for any $\varepsilon > 0$,

$$\begin{aligned}
\lim_{n\to\infty} P\left(|X_n - c| \geq \varepsilon\right) &= \lim_{n\to\infty}\left[P\left(X_n \leq c-\varepsilon\right) + P\left(X_n \geq c+\varepsilon\right)\right] \\
&= \lim_{n\to\infty} P\left(X_n \leq c-\varepsilon\right) + \lim_{n\to\infty} P\left(X_n \geq c+\varepsilon\right) \\
&= 0 + \lim_{n\to\infty} P\left(X_n \geq c+\varepsilon\right) \quad \text{(since } \lim_{n\to\infty} F_{X_n}\left(c-\varepsilon\right)=0) \\
&\leq \lim_{n\to\infty} P\left(X_n > c+\frac{\varepsilon}{2}\right) \\
&= 1\text{-}\lim_{n\to\infty} F_{X_n}\left(c+\frac{\varepsilon}{2}\right) \\
&= 0 \qquad\qquad \text{(since } \lim_{n\to\infty} F_{X_n}\left(c+\frac{\varepsilon}{2}\right)=1).
\end{aligned}$$

Since $\lim_{n\to\infty} P\left(|X_n - c| \geq \varepsilon\right) \geq 0$, we conclude that

$$\lim_{n\to\infty} P\left(|X_n - c| \geq \varepsilon\right) = 0, \text{ for all } \varepsilon > 0,$$

which means $X_n \xrightarrow{p} c.$

Example. Let $X_1, X_2, ...$ be a sequence of i.i.d. $Uniform(0,1)$ random variable. Define the sequence Y_n as

$$Y_n = \min(X_1, X_2, ..., X_n).$$

Prove the following convergence results independently.

a) $Y_n \xrightarrow{d} 0,$ **b)** $Y_n \xrightarrow{p} 0,$ **c)** $Y_n \xrightarrow{L^r} 0,$ for $r \geq 1,$ **d)** $Y_n \xrightarrow{a.s} 0.$

Solution.

a) $Y_n \xrightarrow{d} 0$: Note that

$$F_{X_n}(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

Also, note that $R_{Y_n} = [0,1]$. For $0 \leq y \leq 1$, we can write

$$\begin{aligned}
F_{Y_n}(y) &= P(Y_n \leq y) \\
&= 1 - P(Y_n > y) \\
&= 1 - P(X_1 > y, X_2 > y, ..., X_n > y) \\
&= 1 - P(X_1 > y)P(X_2 > y)\cdots P(X_n > y) \text{(since } X_i\text{'s are independent)} \\
&= 1 - \left(1 - F_{X_1}(y)\right)\left(1 - F_{X_2}(y)\right)\cdots\left(1 - F_{X_n}(y)\right) \\
&= 1 - \left(1 - y\right)^n.
\end{aligned}$$

Therefore, we conclude

$$\lim_{n\to\infty} F_{Y_n}(y) = \begin{cases} 0 & y \leq 0 \\ 1 & y > 0 \end{cases}$$

Therefore, $Y_n \xrightarrow{d} 0.$

b) $Y_n \xrightarrow{p} 0:$ Note that as we found in part a)

$$F_{Y_n}(y) = \begin{cases} 0 & y < 0 \\ 1-(1-y)^n & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

In particular, note that Y_n is a continuous random variable. To show $Y_n \xrightarrow{p} 0,$ we need to show that

$$\lim_{n\to\infty} P\left(|Y_n| \ge \varepsilon\right) = 0, \text{ for all } \varepsilon > 0.$$

Since $Y_n \ge 0,$ it suffices to show that

$$\lim_{n\to\infty} P\left(Y_n \ge \varepsilon\right) = 0, \text{ for all } \varepsilon > 0.$$

For $\varepsilon \in (0,1),$ we have

$$\begin{aligned} P\left(Y_n \ge \varepsilon\right) &= 1 - P\left(Y_n < \varepsilon\right) \\ &= 1 - P\left(Y_n \le \varepsilon\right) \text{ (since } Y_n \text{ is a continuous random variable)} \\ &= 1 - F_{Y_n}(\varepsilon) \\ &= \left(1-\varepsilon\right)^n. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{n\to\infty} P\left(|Yn| \ge \varepsilon\right) &= \lim_{n\to\infty}\left(1-\varepsilon\right)^n \\ &= 0, \text{ for all } \varepsilon \in (0,1]. \end{aligned}$$

c) $Y_n \xrightarrow{L^r} 0,$ for all $r \ge 1:$ By differentiating $F_{Y_n}(y),$ we obtain

$$f_{Y_n}(y) = \begin{cases} n(1-y)^{n-1} & 0 \le y \le 1 \\ 0 & otherwise \end{cases}$$

Thus, for $r \ge 1,$ we can write

$$\begin{aligned} E|Y_n|^r &= \int_0^1 ny^r(1-y)^{n-1}dy \\ &\le \int_0^1 ny(1-y)^{n-1}dy \text{ (since } r \ge 1) \\ &= \left[-y(1-y)^n\right]\Big|_0^1 + \int_0^1 (1-y)^n\, dy \quad \text{(integration by parts)} \\ &= \frac{1}{n+1}. \end{aligned}$$

Therefore,

$$\lim_{n\to\infty}\left(E|Y_n|^r\right)=0.$$

d) $Y_n \xrightarrow{a.s} 0$: We will prove

$$\sum_{n=1}^{\infty} P\left(|Y_n|>\varepsilon\right)<\infty,$$

Which implies $Y_n \xrightarrow{a.s} 0.$ By our discussion in part b),

$$\sum_{n=1}^{\infty} P\left(|Y_n|>\varepsilon\right)=\sum_{n=1}^{\infty}(1-\varepsilon)^n$$

$$=\frac{1-\varepsilon}{\varepsilon}<\infty \quad \text{(geometric series).}$$

4. Conclusions

The point of this paper is to study about the limiting behavior of a sequence of random variables $X_1, X_2, \ldots$ and the relationship between modes of convergences of a sequence of random variables.
Modes of convergences of a sequence of random variables is important concept in probability theory, and its applications to statistics and stochastic processes, for that reason this is very interesting theory to study for.

References

[1]. Feller, W., 1968. An Introduction to Probability Theory and Its Applications. New York: John Wiley and Sons, Vol. 1, 3rd ed.
[2]. Fraser, D. A. S., 1976. Probability and Statistics. Boston: Duxbury Press.
[3]. Hoel, P. G., Port, S., and Stone, C. L., 1971. Introduction to Probability Theory. Boston: Houghton-Mifflin.
[4]. Llukan P., 2008. Probabilitetet dhe Statistika e Zbatuar: Konceptet themelore. Shtëpia Botuese e Libirit Shkollor e Re.
[5]. Morris H. D., Mark J. S., 2012. Probability and statistics (5rd ed.), Pearson Education, Inc, ISBN 10: 0-321-50046-6, ISBN 13: 978-0-321-50046-5
[6]. Patrick B., 1987. Weak convergence of measures: applications in probability. Society for Industrial Mathematics.
[7]. Thomas A. Severini, 2005. Elements of Distribution Theory. Cambridge University Press. USA:New York.