CONVERGENCE IN DISTRIBUTION AND DISTRIBUTIONAL CHARACTERISTIC GROUP

Petar Kopanov

Abstract. *In this work, the relationship between convergence in distribution and convergence in probability of random variables is considered. The notion of a characteristic group of a distribution is introduced to clarify the relationship between these types of convergence.*

Key words: probability distribution, convergence of random variables.

Types of convergence of series of random variables

In probability theory, there are several different concepts of convergence of random variables. The convergence of series of random variables to some limiting random variable is an important concept in probability theory and its applications in statistics and stochastic processes.

In fact, the convergence of series of random variables is a topic of mathematical analysis. The types of convergence of series of random variables can be divided into two large groups:

- The first group concerns only the distributions of the random variables and, therefore, the measures generated by them on the Borel σ-algebra. Traditionally, from this group of convergences in probability theory, only one is studied – that of distribution.
- The second group of convergences is richer and significantly more used. This includes all convergences of random variables or measurable functions on an abstract space with measure.

Convergence of random variables

Definition 1.1. *We say that the sequence of random variables $\{\xi_n\}_{n=1}^{\infty}$ converges to the random variable ξ in distribution if the sequence of their distributions $\{F_n\}_{n=1}^{\infty}$ tends to the distribution F of ξ in every point of continuity of F:*

$$F_{\xi_n}(x) \to F_{\xi}(x), \quad n \to \infty \; \forall x \in C(F_{\xi}).$$

We will denote this convergence as follows:

$$\xi_n \xrightarrow{d} \xi, \quad n \to \infty.$$

The following definition of distributional convergence is also used:

We say that the sequence of random variables $\{\xi_n\}_{n=1}^{\infty}$ *converges to the random variable* ξ *in distribution if* $E(h(\xi_n) \to E(h(\xi)),\ n \to \infty$ *for any continuous function* h.

The definition of distributional convergence by expectation is appropriate in extensions to higher dimensions and to more general function spaces, such as the space $C[0,1]$ of continuous functions on the interval $[0,1]$, provided with a uniform topology.

Definition 1.2. *We say that the sequence of random variables* $\{\xi_n\}_{n=1}^{\infty}$ *converges to the random variable* ξ *in probability iff, for every* $\varepsilon > 0$

$$P\left(|\xi_n - \xi| > \varepsilon\right) \to 0, \quad n \to \infty.$$

Notation:

$$\xi_n \xrightarrow{p} \xi, \quad n \to \infty.$$

Definition 1.3. *We say that the sequence of random variables* $\{\xi_n\}_{n=1}^{\infty}$ *converges to the random variable* ξ *in* r*-mean,* $r > 0$*, iff*

$$E\left|\xi_n - \xi\right|^r \to 0, \quad n \to \infty.$$

Notation:

$$\xi_n \xrightarrow{r} \xi, \quad n \to \infty.$$

When $r = 2$ *this convergence is called convergence in square mean.*

Definition 1.4. *We say that the sequence of random variables* $\{\xi_n\}_{n=1}^{\infty}$ *converges to the random variable* ξ *almost surely (a.s.) iff*

$$P\left(\{\omega : \ |\xi_n(\omega) - \xi(\omega)| \to 0\}\right) = 1, \quad n \to \infty.$$

Notation:

$$\xi_n \xrightarrow{a.s.} \xi, \quad \xi_n \xrightarrow{n.c.} \xi, \quad n \to \infty.$$

We will also consider a less common but very useful concept of convergence, introduced in 1947 by Hsu and Robbins, which is closely related to the Borel-Cantelli lemmas.

Definition 1.5. *We say that the sequence of random variables* $\{\xi_n\}_{n=1}^{\infty}$ *converges completely to the random variable* ξ, *if*

$$\sum_{n=1}^{\infty} P\left(|\xi_n - \xi| > \varepsilon\right) < \infty, \quad \forall\varepsilon > 0.$$

Notation:

$$\xi_n \overset{c.c.}{\to} \xi, \quad \xi_n \overset{н.с.}{\to} \xi, \quad n \to \infty.$$

Relations between different types of convergence of random variables

Between the different types of convergence there are natural connections reflected in Figure 1. The arrows indicate from which of the convergences another follows.

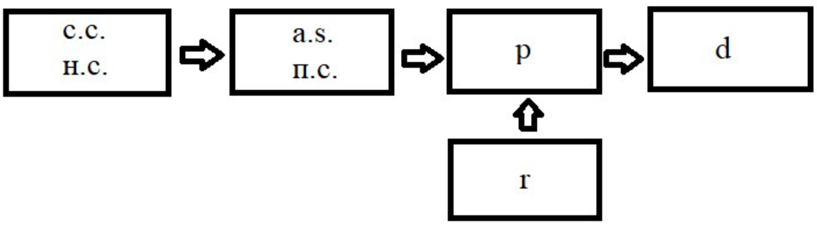

Figure 1. Relations between different types of convergence

It is shown by counterexamples that there are no dependencies between convergence types other than those shown in Figure 1 (see [1, 2]).

Example 1. Convergence in distribution does not entail convergence in probability.

$$\xi_n \overset{d}{\to} \xi \not\Rightarrow \xi_n \overset{p}{\to} \xi.$$

Consider the random variable ξ taking the values 1 and -1 with equal probabilities and the sequence $\{\xi_n\}$, $\xi_n = (-1)^n\xi$. All distributions are the same, but $P\left(|\xi_n - \xi| > \varepsilon\right)$ does not tend to 0.

There is one important exception where convergence in distribution implies convergence in probability: when the limiting random variable is a constant.

Distributional Characteristic Group

We will summarize the idea from Example 1.

Let ξ be a random variable with distribution function $F_\xi(x)$ and $\eta = g(\xi)$ be a new random variable obtained from ξ using the transformation g with distribution function $\xi F_{g(\xi)}(x)$. We consider the following question: for a given distribution $F_\xi(x)$ what are the transformations g for which $F_\xi(x) = F_{g(\xi)}(x)$.

It is not difficult to establish that the set of such transformations $G_F = \{g : \ F_\xi(x) = F_{g(\xi)}(x)\}$ actually represents a group.

We will call this group the distributional characteristic group of the distribution $F_\xi(x)$.

An interesting question directly related to distributional and probability convergences is whether, for the given distribution $F_\xi(x)$ there are other elements in G besides the identity, and if so can they be found explicitly?

An obvious case of such a transformation is $g(\xi) = -\xi$ in case ξ is a symmetric random variable (eg normally distributed with expectation 0).

Examples:

Cauchy distribution and $g(x) = -\frac{1}{x}$.

Binomial distribution: $\xi \sim Bi(n, 0.5)$, $g(k) = n - k$.

The uniform discrete and uniform continuous distributions are the “extreme” examples of distributions whose characteristic group is the group of all permutations.

The presence of elements g other than the identity explains why convergence in distribution does not follow convergence in probability.

Moreover: if the random variable has a distribution $F_\xi(x)$, distributional characteristic group G_F does not consist only of the identity, then we have $\xi_n \xrightarrow{d} \xi \not\Rightarrow \xi_n \xrightarrow{p} \xi$ (from convergence in distribution does not follow convergence in probability).

This is a sufficiently serious reason to pose the problem of finding and describing the characteristic group G_F.

We will briefly consider the two main cases of types of random variables.

1. Discrete random variables

If the values of the distribution are at the points $x_1,\ x_2,\ \ldots,\ x_n$, for the transformation g we can choose an arbitrary permutation of these points, preserving the probabilities in them.

If for this permutation $P\{g(x_i)\} = P\{x_i\}$ holds for all points then g is such a permutation.

Not every discrete distribution has this property, but some do. Obviously, for the discrete uniform distribution all permutations are such transformations. Another example is the binomial distribution with $p = 1/2$. Then the permutation $g(k) = n - k$ satisfies the condition.

The general scheme is as follows: the points are divided into groups $X_1 = \{x_{11},\ x_{12},\ \ldots,\ x_{1n1}\}$, $X_2 = \{x_{21},\ x_{22},\ \ldots,\ x_{2n2}\}$, $\ldots$, $X_k = \{x_{k1},\ x_{k2},\ \ldots,\ x_{knk}\}$, as the points in each group are equally likely: $p(x_{js}) = p_j$, $n_1.p_1 + n_2.p_2 + \cdots + n_k.p_k = 1$. In this case, the characteristic group $G_F = S_{n1}\mathrm{x}S_{n2}\mathrm{x}\ldots\mathrm{x}S_{nk}$ is a Cartesian product of the permutation groups of the corresponding subsets.

The distributional characteristic group G_F is trivial if and only if there are no 2 distinct points whose probabilities are equal.

2. Absolutely continuous random variables

Analogously to the case of discrete distributions, if g is a transformation for which $f(g(x)) = f(x)$, where $f(x)$ is the density of the distribution, we will have $F_\xi(x) = F_{g(\xi)}(x)$.

The essential difference in this case is that we may have infinitely many subsets of points in which the density is the same.

Below we give some geometric examples of densities with points of equal density values:

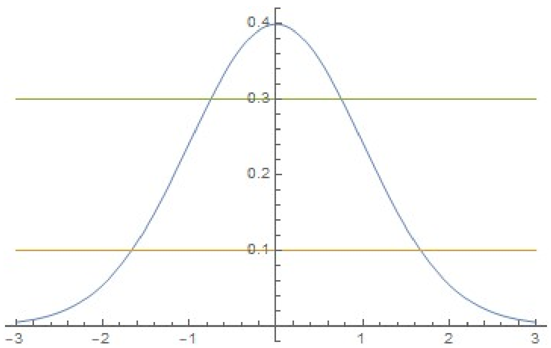

Figure 2. Normal distribution

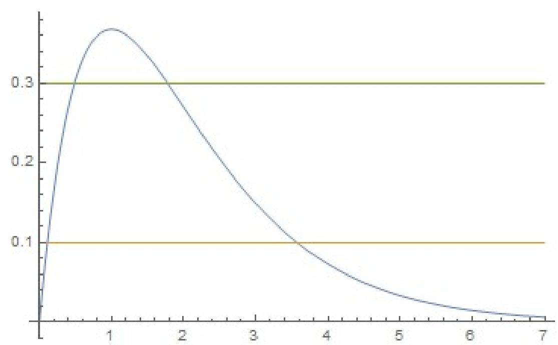

Figure 3. Gamma distribution

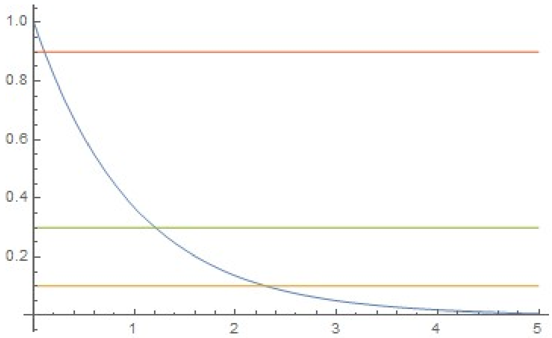

Figure 4. Exponential distribution

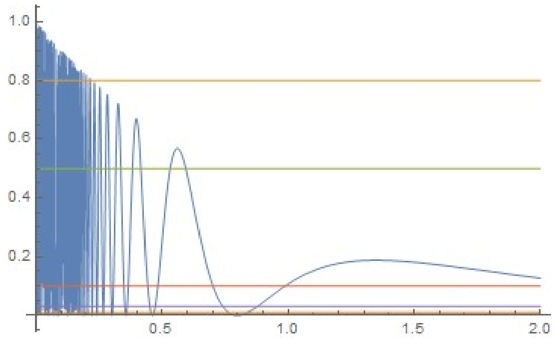

Figure 5. A distribution with infinitely many points with equal densities: $f(x) = e^{-x}.\cos(1/x^2)^2$

From the above reasoning, it follows that a necessary condition to have convergence in probability of a series of random variables if we have convergence in distribution of this series is that the characteristic group of the marginal distribution is trivial: $g = \{e\}$.

We will conclude with a question: is this condition a sufficient condition?

References

[1] J. Stoyanov, *Counterexamples in Probability*, Dover 2014, ISBN-13: 978-0486499987.

[2] A. Gut, *Probability: a graduate course*, Springer, 2013, ISBN: 978-1-4614-4708-5.

Petar Kopanov*
"Paisii Hilendarski" University of Plovdiv,
Faculty of Mathematics and Informatics,
236 Bulgaria Blvd., 4003 Plovdiv, Bulgaria
Corresponding author: pkopanov@uni-plovdiv.bg