# Frequently Asked Questions

1. What do you mean by Convergence of random variables?

**Answer:**

The convergence of sequences of random variables to some limit random variable is an important concept in probability theory. It has a vital role to play in statistics and in stochastic processes. In probability theory, there exist several different notions of **convergence of random variables**.

Convergence of random variables formalize the idea that a sequence of essentially random or unpredictable events can sometimes be expected to settle down into a behaviour that is essentially unchanging when items far enough into the sequence are studied.

The different possible notions of convergence relate to how such behaviour can be characterised: two readily understood behaviours are that the sequence eventually takes a constant value and that values in the sequence continue to change but can be described by an unchanging probability distribution.

2. Define convergence in distribution.

**Answer:**

A sequence $\{X_1, X_2, \ldots\}$ of random variables is said to **converge in distribution**, or **converge weakly**, or **converge in law** to a random variable $X$ if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

For every number $x$ in R at which $F$ is continuous. Here $F_n$ and $F$ are the cumulative distribution functions of random variables $X_n$ and X correspondingly.

3. Why we require that the distribution functions converge only at continuity points for the limiting distribution function?

**Answer:**

The requirement that only the continuity points of $F$ should be considered is essential. An example explains why we require that the distribution functions converge only at continuity points for the limiting distribution function

If $X_n$ are distributed uniformly on intervals $[0, {}^1\!/_n]$, then this sequence converges in distribution to a degenerate random variable $X = 0$. Indeed, $F_n(x) = 0$ for all $n$ when $x \le 0$, and $F_n(x) = 1$ for all $x \ge {}^1\!/_n$ when $n > 0$.

However, for this limiting random variable, $F(0) = 1$, even though $F_n(0) = 0$ for all $n$. Thus the convergence of cumulative distribution functions fails at the point $x =$

0 where $F$ is discontinuous.

4. What is the advantage of convergence in distribution as compared to other modes of convergence?

**Answer:**

The first, obvious fact to notice is that convergence in distribution only involves the distributions of the random variables. Thus, the random variables need not even be defined on the same probability space (that is, they need not be defined for the same random experiment). This is in sharp contrast and advantage to the other modes of convergence.

5. How do we denote convergence in distribution and what we mean by it?

**Answer:**

Convergence in distribution may be denoted as

$$Xn \xrightarrow{d} X, Xn \xrightarrow{D} X, Xn \xrightarrow{L} X,$$

For example, if $X$ is standard normal we can write $X_n \xrightarrow{d} \mathcal{N}(0, 1)$.

It should be clear what we mean by $Xn \xrightarrow{d} X$,

The random variables Xn converge in distribution to a random variable X having distribution function F. Similarly, we have $Fn \xrightarrow{d} F$

if there is a sequence of random variables {Xn}, where Xn has distribution function Fn, and a random variable X having distribution function F, so that $Xn \xrightarrow{d} X$,

6. What do you mean by weak convergence?

**Answer:**

A sequence of random elements $\{X_n\}$ converges weakly to $X$ if

$$\mathrm{E}^* h(X_n) \to \mathrm{E}\, h(X)$$

For all continuous bounded functions $h(\cdot)$. Here E* denotes the *outer expectation*, that is the expectation of a "smallest measurable function $g$ that dominates $h(X_n)$".

7. What are the properties of convergence in distribution?

**Answer:**

- Since $F(a) = P(X \leq a)$, the convergence in distribution means that the probability for $X_n$ to be in a given range is approximately equal to the probability that the value of $X$ is in that range, provided $n$ is sufficiently large
- In general, convergence in distribution does not imply that the sequence of corresponding probability density functions will also converge

- As an example one may consider random variables with densities $f_n(x) = (1 - \cos(2\pi nx))$ These random variables converge in distribution to a uniform $U(0,1)$, whereas their densities do not converge at all
- Portmanteau lemma provides several equivalent definitions of convergence in distribution. Although these definitions are less intuitive, they are used to prove a number of statistical theorems
- Continuous mapping theorem states that for a continuous function $g(\cdot)$, if the sequence $\{X_n\}$ converges in distribution to $X$, then so does $\{g(X_n)\}$ converge in distribution to $g(X)$
- Lévy's continuity theorem: the sequence $\{X_n\}$ converges in distribution to $X$ if and only if the sequence of corresponding characteristic functions $\{\varphi_n\}$ converges point wise to the characteristic function $\varphi$ of $X$
- Convergence in distribution is metrizable by the Lévy–Prokhorov metric
- A natural link to convergence in distribution is the Skorokhod's representation theorem

8. Define convergence in distribution in terms of Moment Generating function.

**Answer:**

The sequence $\{X_n\}$ converges in distribution to $X$ if and only if the sequence of corresponding moment generating functions $M_{X_n}(t)$ converges to the moment generating function $M_X(t)$ of $X$. That is if

$$M_{X_n}(t) \longrightarrow M_X(t) \text{ then } Xn \xrightarrow{d} X$$

9. Explain convergence in distribution with examples.

**Answer:**

Examples:

**Dice factory**

Suppose a new dice factory has just been built. The first few dice come out quite biased, due to imperfections in the production process. The outcome from tossing any of them will follow a distribution markedly different from the desired uniform distribution

As the factory is improved, the dice become less and less loaded, and the outcomes from tossing a newly produced dice will follow the uniform distribution more and more closely.

**Tossing coins**

Let Xn be the fraction of heads after tossing up an unbiased coin n times. Then X1 has the Bernoulli distribution with expected value μ = 0.5 and variance σ2 = 0.25. The subsequent random variables X2, X3 … will all be distributed binomially.

As n grows larger, this distribution will gradually start to take shape more and more similar to the bell curve of the normal distribution. If we shift and rescale Xn's appropriately, Then, $Zn = \frac{\sqrt{n}(\overline{X}n - \mu)}{\sigma}$ will be converging in distribution to the standard normal, the result that follows from the celebrated central limit theorem

10. What do you mean by convergence in probability?

**Answer:**

The concept of convergence in probability is based on the following intuition: two random variables are "close to each other" if there is a high probability that their difference is very small. The basic idea behind this type of convergence is that the probability of an “unusual” outcome becomes smaller and smaller as the sequence progresses.

A sequence $\{X_n\}$ of random variables converges in probability towards X if for all ε > 0

$$\lim_{n\to\infty} \Pr\left(|X_n - X| \geq \varepsilon\right) = 0.$$

11. What are the properties of convergence in probability?

**Answer:**
Properties of convergence in probability are

- Convergence in probability implies convergence in distribution.
- Convergence in probability does not imply almost sure convergence.
- In the opposite direction, convergence in distribution implies convergence in probability only when the limiting random variable *X* is a constant.
- The continuous mapping theorem states that for every continuous function *g*(·), if $X_n \xrightarrow{p} X$, then also $g(X_n) \xrightarrow{p} g(X)$.
- Convergence in probability defines a topology on the space of random variables over a fixed probability space.

12. Briefly explain convergence in probability using examples.

**Answer:**

**Height of a person:**

Consider the following experiment. First, pick a random person in the street. Let X be his/her height, which is ex antea random variable. Then you start asking other people to estimate this height by eye. Let Xn be the average of the first n responses. Then (provided there is no systematic error) by the law of large numbers, the sequence Xn will converge in probability to the random variable X.

**Archer:**

Suppose a person takes a bow and starts shooting arrows at a target. Let Xn be his score in n-th shot. Initially he will be very likely to score zeros, but as the time goes and his archery skill increases, he will become more and more likely to hit the bullseye and score the maximum of 10 points. After the years of practice the probability that he hit anything but 10 will be getting increasingly smaller and smaller. Thus, the sequence Xn converges in probability to X=10.

13. Show that if $Xn \xrightarrow{P} X$ then, $Xn \xrightarrow{d} X$ .

**Answer:**

P(Xn ≤ t) = P({Xn ≤ t} ∩ {|Xn − X| ≤ e}) + P({Xn ≤ t} ∩ {|Xn − X| > e})

≤ P(X ≤ t + e) + P(|Xn − X| > e)

First pick e small enough so that P(X ≤ t + e) ≤ P(X ≤ t) + η/2. (Since F is right continuous.) Then pick n large enough so that P(|Xn − X| > e) < η/2. Then, for n large enough, we have

$$P[Xn \le t] \le P[X \le t+e] + P[\left|Xn - X\right| > e] \Rightarrow Fn(t) \le P[X \le t) + \eta/2 + \eta/2$$

$$= F(t) + \eta$$

Therefore,

Fn(t) ≤ F(t) + η .

Similarly, for n large enough, if F is continuous at t, we have

Fn(t) ≥ F(t)−η. This shows that

$\lim_{n \to \infty} Fn(t) = F(t)$ at continuity points of F.

Hence, $Xn \xrightarrow{P} X$ implies $Xn \xrightarrow{d} X$

14. Whether the converse of the above (Q.No.13) is true?

**Answer:**

The converse of the above is not true, that is $Xn \xrightarrow{d} X$ does not imply $Xn \xrightarrow{P} X$
But as $Xn \xrightarrow{d} C$ then $Xn \xrightarrow{P} C$ where C is a constant.

Thus, when the limit is a constant, convergence in probability and convergence in distribution are equivalent.

15. What are the applications of convergence of probability?

**Answer:**

The basic idea behind this type of convergence is that the probability of an “unusual” outcome becomes smaller and smaller as the sequence progresses. The concept of convergence in probability is used very often in statistics. For example, an estimator is called consistent if it converges in probability to the quantity being estimated. Convergence in probability is also the type of convergence established by the weak law of large numbers. Hence, this mode of convergence can be used to obtain properties of estimators as the sample a size tends to infinity.