

Convergence Concepts of Random Variables and Functions

©2002–2007, Professor Seppo Pynnönen,
Department of Mathematics and Statistics,
University of Vaasa

Version: January 5, 2007

Convergence Modes

Convergence of real numbers and real valued functions

(a) Limit of a sequence of real numbers:

Definition 1 The real number $a \in \mathbb{R}$ is the limit of a sequence $\{x_n\}$, denoted as

$$\lim_{n \rightarrow \infty} x_n = a$$

if for any $\epsilon > 0$ there is an integer $n(\epsilon)$ such that

$$|x_n - a| < \epsilon \text{ whenever } n \geq n(\epsilon),$$

where $n(\epsilon)$ indicates that it may depend on ϵ .

(b) Limit of a sequence of functions:

Definition 2 A sequence of real-valued functions $\{f_n\}$ defined on a set $S \subset \mathbb{R}$ converge point wise to a function $f : S \rightarrow \mathbb{R}$, denoted as $\lim f_n(x) = f(x)$, if for any $\epsilon > 0$ and $x \in S$ there is an integer $n(\epsilon, x)$ such that

$$|f_n(x) - f(x)| < \epsilon \text{ for each } x \in S, \text{ whenever } n \geq n(\epsilon, x),$$

where $n(\epsilon, x)$ indicates that the integer may depend on ϵ and x .

If $n(\epsilon, x) = n(\epsilon)$, and for any $\epsilon > 0$

$$|f_n(x) - f(x)| < \epsilon \text{ for all } x \in S, \text{ whenever } n \geq n(\epsilon),$$

then $\{f_n\}$ is converging uniformly to f .

Note. In the latter case whenever $n \geq n(\epsilon)$ the graphs of $f_n(x)$ become indistinguishable from $f(x)$.

Note. The uniform convergence implies the pointwise convergence, but not vice versa.

Example. Let $f_n(x) = x^n$ if $0 \leq x \leq 1$. Then

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{if } 0 \leq x < 1, \\ 1 & \text{if } x = 1. \end{cases}$$

Thus $\{f_n\}$ converge pointwise but not uniformly to the limit function $f(x) = 0$ for $0 \leq x < 1$ and $f(x) = 1$ for $x = 1$.

Definition 3 (Limit of a function) A point c is the limit of a function $f(x)$ as $x \rightarrow p$

$$\lim_{x \rightarrow p} f(x) = c$$

if for any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - c| < \epsilon, \text{ whenever } 0 < |x - p| < \delta.$$

Note. f need not be defined at p .

Example. It is well known that

$$f(x) = \frac{\sin x}{x} \rightarrow 1 \text{ as } x \rightarrow 0.$$

Definition 4 A function $f : S \rightarrow \mathbb{R}$ is continuous at a point $p \in S$ if f is defined at p , and for all $\epsilon > 0$ there exist a $\delta > 0$ such that

$$|f(x) - f(p)| < \epsilon, \text{ whenever } |x - p| < \delta$$

Properties of **lim**:

Theorem 1 Let $\lim x_n = a$ and $f(x)$ is continuous at a then

$$\lim f(x_n) = f(\lim x_n) = f(a).$$

Proof. For an arbitrary $\epsilon > 0$ select $\delta > 0$ (which may depend on ϵ) such that

$$|f(x) - f(a)| < \epsilon, \text{ whenever } |x - a| < \delta.$$

Because $x_n \rightarrow a$ then there exists an n_δ such that $|x_n - a| < \delta$ when $n \geq n_\delta$, and hence

$$|f(x_n) - f(a)| < \epsilon$$

i.e., $\lim_{n \rightarrow \infty} f(x_n) = f(a)$, and the proof is complete.

Examples: If $\lim_{n \rightarrow \infty} x_n = a \neq 0$ then

$$\lim \frac{1}{x_n} = \frac{1}{\lim x_n} = \frac{1}{a}$$

$$\lim |x_n| = |\lim x_n| = |a|.$$

Given that $\lim x_n$ and $\lim y_n$ exist then for an $b, c \in \mathbb{R}$

$$\lim (bx_n + cy_n) = b \lim x_n + c \lim y_n$$

$$\lim x_n y_n = (\lim x_n)(\lim y_n)$$

$$\lim \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n}$$

provided that $\lim y_n \neq 0$.

Convergence of random variables

Suppose next that the sequence X_n is a sequence of random variables defined on a probability space (Ω, \mathcal{F}, P) . Thus $X_n : \Omega \rightarrow \mathbb{R}$ are just real valued functions with domain Ω and values $X_n(\omega) \in \mathbb{R}$, and all the convergence concepts defined so far apply as such. These are, however, pretty *uninteresting*, because they are tied to particular outcome ω . This would be just equivalent to investigate sequences $\{f_n(x)\}$ of real valued functions, and we would lose the new component—*randomness*.

Introducing randomness, then the interest in the convergence modes is the probability of those events where the sequence does not converge. If this diminishes to zero it is reasonable to say that a sequence of random variables converges.

Definition 5 (Convergence with probability one, w.p.1-convergence or almost sure (a.s.) convergence). *A sequence $\{X_n\}$ of random variables converges almost surely to a random variable X , denoted as*

$$X_n \xrightarrow{\text{a.s.}} X \text{ (or } X_n \rightarrow X \text{ w.p.1)}$$

if for any $\epsilon > 0$

$$\lim_{k \rightarrow \infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \epsilon, \text{ whenever } n \geq k\}) = 1.$$

That is

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

Note. The sets $S_k = \{\omega \in \Omega : |X_n - X| < \epsilon \text{ whenever } n \geq k\}$ are an increasing sequence of sets, i.e., $S_k \subset S_{k+1} \subset \dots$, and in set notations can be written as

$$\begin{aligned} S_k &= \{\omega \in \Omega : |X_k(\omega) - X(\omega)| < \epsilon \text{ and } |X_{k+1}(\omega) - X(\omega)| < \epsilon \text{ and } \dots\} \\ &= \bigcap_{n=k}^{\infty} \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \epsilon\}. \end{aligned}$$

Thus $\lim_{k \rightarrow \infty} P(S_k) = 1$ and hence $\lim_{k \rightarrow \infty} P(S_k^c) = 0$, i.e., for the sets $S_k^c = \Omega \setminus S_k = \bigcup_{n=k}^{\infty} \{\omega \in \Omega : |X_n - X| \geq \epsilon\}$, where the convergence does not occur, the probability diminishes to zero.

Note. Confirming almost sure convergence of a sequence is, unfortunately, not an easy exercise. There are many sufficient conditions. One is the so

called

Borel-Cantelli Lemma: Let $\{X_n\}$ be a sequence of random variables. If for every $\epsilon > 0$

$$\sum_{n=1}^{\infty} P(|X_n - X| \geq \epsilon) < \infty$$

then

$$X_n \xrightarrow{a.s.} X.$$

Proof. Now $P(|X_n - X| \geq \epsilon) \geq 0$, and by assumption the sum above converges, which together imply that

$$\lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} P(|X_n - X| \geq \epsilon) = 0.$$

But then

$$\begin{aligned} \lim_{k \rightarrow \infty} P(\cap_{n=k}^{\infty} \{|X_n - X| < \epsilon\}) &= 1 - \lim_{k \rightarrow \infty} P(\cup_{n=k}^{\infty} \{|X_n - X| \geq \epsilon\}) \\ &\geq 1 - \lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} P(|X_n - X| \geq \epsilon) = 1, \end{aligned}$$

and the proof is complete.

Definition 6 A sequence of random variables $\{X_n\}$ converges in probability to random variable X , denoted as

$$\text{plim } X_n = X \quad \text{or} \quad X_n \xrightarrow{P} X$$

if for all $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \epsilon) = 0.$$

Assume $X_n \xrightarrow{a.s.} X$. Obviously, because $\{|X_n - X| \geq \epsilon\} \subset \cup_{k=n}^{\infty} \{|X_k - X| \geq \epsilon\}$, so that

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \epsilon) \leq \lim_{n \rightarrow \infty} P(\cup_{k=n}^{\infty} \{|X_k - X| \geq \epsilon\}) = 0.$$

Thus a.s. convergence implies convergence in probability.

The converse does not hold.

Example. Let $\{X_n\}$ be a sequence of independent random variables, such that

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

Then for any $1 > \epsilon > 0$

$$P(|X_n| < \epsilon) = P(X_n = 0) = 1 - \frac{1}{n} \rightarrow 1, \text{ as } n \rightarrow \infty.$$

So $X_n \xrightarrow{P} 0$. Nevertheless,

$$\begin{aligned} \lim_{k \rightarrow \infty} P(\cup_{n=k}^{\infty} \{|X_n| \geq \epsilon\}) &= 1 - \lim_{k \rightarrow \infty} P(\cap_{n=k}^{\infty} \{|X_n| < \epsilon\}) \\ &= 1 - \lim_{k \rightarrow \infty} \prod_{n=k}^{\infty} P(|X_n| < \epsilon) \quad (\text{by independence}) \\ &= 1 - \lim_{k \rightarrow \infty} \prod_{n=k}^{\infty} (1 - \frac{1}{n}) = 1 \end{aligned}$$

since $\prod_{n=k}^{\infty} (1 - \frac{1}{n}) = 0$ for all k .¹

Thus $X_n \rightarrow 0$ in probability but not a.s.

Definition 7 (Convergence in distribution) *A sequence of random variables $\{X_n\}$ in distribution to a random X , denoted as*

$$X_n \xrightarrow{D} X$$

if

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

in all continuity points x of $F(x)$, where $F_n(x) = P(X_n \leq x)$ is the cumulative distribution function (cdf) of X_n and $F(x) = P(X \leq x)$ is the cdf of X .

If $g(\cdot)$ is a continuous monotonically increasing function. Then if $Y_n = g(X_n)$, we have

$$F_{Y_n}(y) = P(Y_n \leq y) = P(g(X_n) \leq y) = P(X_n \leq g^{-1}(y)) = F_{X_n}(g^{-1}(y)).$$

Similarly, if $Y = g(X)$ then

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

¹ $\prod_{m=2}^n (1 - \frac{1}{m}) = \prod_{m=2}^n (\frac{m-1}{m}) = \frac{(n-1)!}{n!} = \frac{1}{n} \rightarrow 0$ as $n \rightarrow \infty$. Generally $\prod_{m=k}^n (1 - \frac{1}{m}) = (\prod_{m=2}^{k-1} (1 - \frac{1}{m}))^{-1} \prod_{m=2}^n (1 - \frac{1}{m}) = (\frac{1}{k-1})^{-1} \frac{1}{n} = (k-1) \frac{1}{n} \rightarrow 0$ as $n \rightarrow \infty$.

If $X_n \xrightarrow{D} X$, then

$$F_{Y_n}(y) = F_{X_n}(g^{-1}(y)) \rightarrow F_X(g^{-1}(y)) = F_Y(y) \text{ as } n \rightarrow \infty.$$

That is $g(X_n) \xrightarrow{D} g(X)$.

More generally:

Theorem 2 (Continuous Mapping Theorem) *If $X_n \xrightarrow{D} X$ and $g(\cdot)$ is continuous, then $g(X_n) \xrightarrow{D} g(X)$.*

Definition 8 (Convergence in quadratic mean) *A sequence of random variables $\{X_n\}$ with $E(X_n^2) < \infty$ converges in quadratic mean to X (with $E(X^2) < \infty$), denoted as*

$$X_n \xrightarrow{q.m.} X$$

if

$$\lim_{n \rightarrow \infty} E(X_n - X)^2 = 0.$$

Using the Chebysev inequality we obtain directly that if $X_n \xrightarrow{q.m.} X$ then $X_n \xrightarrow{P} X$. Generally

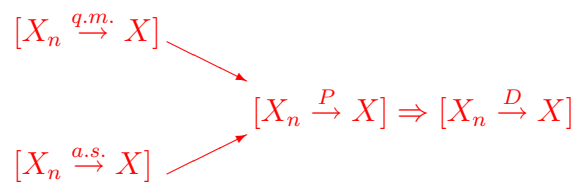


Figure. Relations between convergence modes.

Furthermore

Theorem 3 *Let $\{X_n\}$ and $\{Y_n\}$ be sequences of random variables.*

(a) *If c is a constant then $X_n \xrightarrow{q.m.} c \iff X_n \xrightarrow{P} c$*

(b) If c is constant then $X_n \xrightarrow{q.m.} c \iff \lim E[X_n] = c$ and $\lim \text{Var}[X_n] = \lim E(X_n - E[X_n])^2 = 0$.

(c) If $\text{plim } X_n$ and $\text{plim } Y_n$ exist and a and b are constants, then

(1) $\text{plim } (aX_n \pm bY_n) = a \text{plim } X_n \pm b \text{plim } Y_n$,

(2) $\text{plim } (X_n Y_n) = (\text{plim } X_n)(\text{plim } Y_n)$

(3) if $\text{plim } Y_n \neq 0$ then $\text{plim } \frac{X_n}{Y_n} = \frac{\text{plim } X_n}{\text{plim } Y_n}$.

(4) if $X_n \xrightarrow{D} X$ and $\text{plim } X_n = \text{plim } Y_n$, then $Y_n \xrightarrow{D} X$

(d) If $X_n \xrightarrow{D} X$ and $\text{plim } Y_n = c$, a constant, then

(1) $X_n \pm Y_n \xrightarrow{D} X \pm c$.

(2) $X_n Y_n \xrightarrow{D} cX$ given that $c \neq 0$.

(3) If $c = 0$, then $\text{plim } X_n Y_n = 0$.

We observe that "plim" shares the same properties as the usual \lim , but the other stochastic convergence concepts do not.

Especially (and unfortunately) q.m. limit shares very few of these properties. For example $X_n \xrightarrow{q.m.} X$ and $Y_n \xrightarrow{q.m.} Y$ does not imply $X_n + Y_n \xrightarrow{q.m.} X + Y$! (If it did, then by defining $Z_n = X_n - Y_n$ one gets $Z_n \rightarrow 0$, and property (c) would make q.m. equivalent to plim.)