Modes of Convergence

December 2, 2008

Definition 1. *Let $X, X_1, X_2, \cdots$ be a sequence of random variables.*

1. *We say that X_n* **converges to** *X* **almost surely** *($X_n \to^{a.s.} X$) if*

$$P\{\lim_{n\to\infty} X_n = X\} = 1.$$

2. *We say that X_n* **converges to** *X* **in** *L^p, $p > 0$, ($X_n \to^{L^p} X$) if,*

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

3. *We say that X_n* **converges to** *X* **in probability** *($X_n \to^{P} X$) if, for every $\epsilon > 0$,*

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

4. *We say that X_n* **converges to** *X* **in distribution** *($X_n \to^{\mathcal{D}} X$) if, for every bounded continuous function $h : \mathbb{R} \to \mathbb{R}$.*

$$\lim_{n\to\infty} Eh(X_n) = Eh(X).$$

For almost sure convergence, convergence in probability and convergence in distribution, if X_n converges to X and if g is a continuous then $g(X_n)$ converges to $g(X)$.

Convergence in distribution differs from the other modes of convergence in that it is based not on a direct comparison of the random variables X_n with X but rather on a comparision of the distributions $P\{X_n \in A\}$ and $P\{X \in A\}$. Using the change of variables formula, convergence in distribution can be written

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} h(x)\, dF_{X_n}(x) = \int_{-\infty}^{\infty} h(x)\ dF_X(x).$$

We can use this to show that $X_n \to^{\mathcal{D}} X$ if and only if

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

for all points x that are continuity points of F_x.

Example 2. *Let X_n be uniform on the points $\{1/n, 2/n, \cdots n/n = 1\}$. Then, using the convergence of a Riemann sum to a Riemann integral, we have as $n \to \infty$,*

$$Eh(X_n) = \sum_{i=1}^{n} h\left(\frac{i}{n}\right)\frac{i}{n} \to \int_0^1 h(x)\,dx = Eh(X)$$

where X is a uniform random variable on the interval $[0, 1]$.

Example 3. *Let $X_i, i \geq 1$, be independent uniform random variable in the interval $[0, 1]$. and let $Y_n = n(1 - X_{(n)})$. Then,*

$$F_{Y_n}(y) = P\{n(1 - X_{(n)} \leq y\} = P\left\{1 - \frac{y}{n} \leq X_{(n)}\right\} = 1 - \left(1 - \frac{y}{n}\right)^n \to 1 - e^{-y}.$$

Thus, the magnified gap between the highest order statistic and 1 converges in distribution to an exponentail random variable.

1 Relationships among the Modes of Convergence

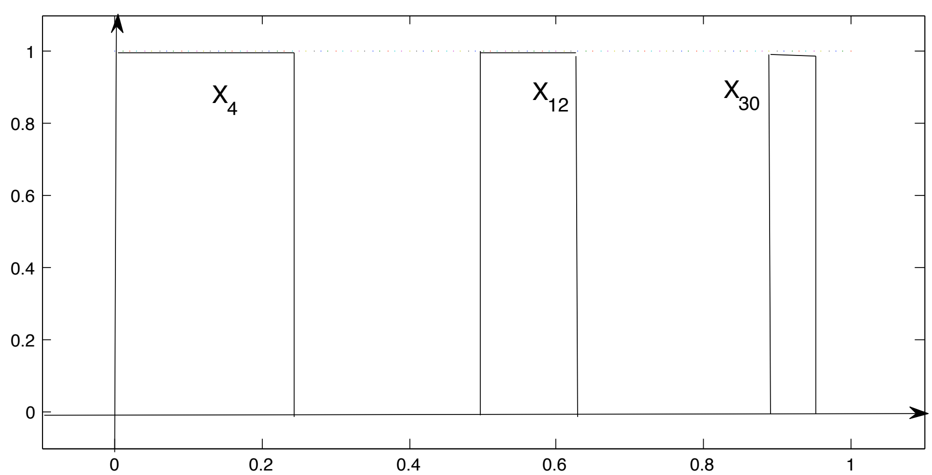
- By examining Chebychev's inequality,

$$P\{|X_n - X| > \epsilon\} \leq \frac{E[|X_n - X|^p]}{\epsilon^p}$$

  we see that convergence in L^p implies convergence in probability.

- Also, convergence almost surely implies convergence in probability.
- Convergence in probability implies convergence in distribution.
- We can write any positive integer by $k + 2^n$, $k = 0, 1, \ldots, 2^{n-1} - 1$ and define

$$X_{k+2^n}(s) = I_{((k-1)2^{-n}, k2^{-n}]}(s), \quad 0 \leq s \leq 1.$$



Then, if our probability space $[0,1]$ has a probability that assigns its length to an interval,

$$P\{|X_{k+2^n}| > \epsilon\} = 2^{-n}$$

and the sequence converges to 0 in probability. However, for each $s \in (0,1]$, $X_j(s) = 1$ and $X_j(s) = 0$ for infinitely many j and so the sequence does not converge almost surely.

- $E|X_{k+2^n} - 0|^p = 2^{-np}$, so the sequence converges in L^p to 0.
- If
$$Y_{k+2^n}(s) = 2^n I_{((k-1)2^{-n}, k2^{-n}]}(s), \quad 0 \le s \le 1.$$
Then, again,
$$P\{|Y_{k+2^n}| > \epsilon\} = 2^{-n}$$
and the sequence converges to 0 in probability. However,
- $E|Y_{k+2^n} - 0| = 2^n P\{Y_{k+2^n} = 2^n\} = 2^n 2^{-n} = 1$, so the sequence does not converge in L^1.

2 Laws of Large Numbers

The best convergences theorem showing that the sample mean converges to the mean of the common distribution is the **strong law of large numbers**

Theorem 4. *Let $X_1, X_2, \ldots$ be independent identically distributed random variables and set $S_n = X_1 + \cdots + X_n$, then*

$$\lim_{n\to\infty} \frac{1}{n} S_n$$

exists almost surely if and only if $E|X_1| < \infty$. In this case the limit is $EX_1 = \mu$ with probability 1.

Convergence laws using a mode other than almost sure is called **a weak law**. Here is a L^2-weak law of large numbers.

Theorem 5. *Assume that $X_1, X_2, \ldots$ for a sequence of real-valued uncorrelated random variable with common mean μ. Futher assume that their variances are bounded by some constant C. Write*

$$S_n = X_1 + \cdots + X_n.$$

Then

$$\frac{1}{n} S_n \to^{L^2} \mu.$$

Proof. Note that $E[S_n/n] = \mu$. Then

$$E[(\frac{1}{n}S_n - \mu)^2] = \mathrm{Var}(\frac{1}{n}S_n) = \frac{1}{n^2}(\mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)) \le \frac{1}{n^2}Cn.$$

Now, let $n \to \infty$ □

Because L^2 convergence implies convergence in probability, we have, in addition,

$$\frac{1}{n} S_n \to^{P} \mu.$$