Convergence of sequences of random variables

February 28, 2018

Convergence of sequences of random variables

The weak law of large numbers

Theorem

Let $X_1, X_2, \ldots, X_k, \ldots$ be a sequence of independent and identically distributed random variables with finite expectation m and finite variance.

Set

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}, \quad n \geq 1.$$

Then, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\overline{X}_n - m\right| \geq \epsilon\right) = 0$$

The weak law of large numbers

The random variable $\overline{X}_n$ has mean m:

$$\mathrm{E}\left(\overline{X}_n\right) = \mathrm{E}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= \frac{\mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n)}{n} = \frac{nm}{n} = m$$

The weak law of large numbers

If $\mathrm{Var}(X_k) = \sigma^2$, then the variance of $\overline{X}_n$ is σ^2/n.

$$\mathrm{Var}\left(\overline{X}_n\right) = \mathrm{E}\left(\left(\overline{X}_n - m\right)^2\right) = \frac{1}{n^2}\,\mathrm{E}\left(\left(\sum_{k=1}^{n}(X_k - m)\right)^2\right)$$

$$= \frac{1}{n^2}\left(\sum_{k=1}^{n}\mathrm{E}\left((X_k - m)^2\right) + 2\sum_{1\le i<j\le n}\underbrace{\mathrm{E}\left((X_i - m)(X_j - m)\right)}_{0}\right)$$

$$= \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{E}\left((X_k - m)^2\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

The weak law of large numbers

Applying Chebyshev's inequality,

$$\mathrm{P}\left(\left|\overline{X}_n - m\right| \ge \epsilon\right) \le \frac{\sigma^2/n}{\epsilon^2}$$

Therefore

$$\mathrm{P}\left(\left|\overline{X}_n - m\right| \ge \epsilon\right) \to 0 \quad \text{as} \quad n \to \infty$$

Probability as the limit of the relative frequency

Let A be an event with probability $\mathrm{P}(A)$.

In each of n independent repetitions of the random experiment, we observe weather or not A occurs. More precisely, for $1 \le k \le n$, let X_k be the indicator random variable of the event "A happens in the k-th repetition".

Hence, $\mathrm{E}(X_k) = \mathrm{P}(A)$.

Moreover,

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = f_n(A)$$

is the relative frequency of the event A.

Probability as the limit of the relative frequency

Therefore, for all $\epsilon > 0$,

$$\mathrm{P}\left(|f_n(A) - \mathrm{P}(A)| \ge \epsilon\right) \to 0 \quad \text{as} \quad n \to \infty$$

Equivalently,

$$\mathrm{P}\left(|f_n(A) - \mathrm{P}(A)| < \epsilon\right) \to 1 \quad \text{as} \quad n \to \infty$$

In a certain sense, the relative frequency of A converges to its probability $\mathrm{P}(A)$.

Convergence in probability

Definition

The sequence $X_1, X_2, \ldots, X_n, \ldots$ converges in probability to the random variable X as $n \to \infty$ if, for all $\epsilon > 0$,

$$\mathrm{P}\left(|X_n - X| \geq \epsilon\right) \to 0 \quad \text{as} \quad n \to \infty$$

Notation:

$$X_n \xrightarrow{\mathrm{P}} X$$

Convergence in probability

- The sequence of sample means $\overline{X_n}$ converges in probability to the common expected value m:

$$\overline{X}_n \xrightarrow{\mathrm{P}} m$$

- The relative frequency $f_n(A)$ converges in probability to $\mathrm{P}(A)$:

$$f_n(A) \xrightarrow{\mathrm{P}} \mathrm{P}(A)$$

Example

Let $X_1, X_2, \ldots, X_n, \ldots$ be random variables such that

$$\mathrm{P}(X_n = 0) = 1 - \frac{1}{n}, \quad \mathrm{P}(X_n = -1) = \mathrm{P}(X_n = 1) = \frac{1}{2n}, \quad n \geq 1$$

Let us prove that

$$X_n \xrightarrow{\mathrm{P}} 0$$

Example

We have to prove that

$$\mathrm{P}(|X_n| < \epsilon) \longrightarrow 1 \quad \text{as} \quad n \to \infty$$

- If $\epsilon > 1$, then $\mathrm{P}(|X_n| < \epsilon) = 1$ for all $n \geq 1$.
- Otherwise, if $\epsilon \leq 1$, then

$$\mathrm{P}(|X_n| < \epsilon) = \mathrm{P}(X_n = 0) = 1 - \frac{1}{n} \to 1$$

The characteristic function of $\overline{X}_n$

Let $M_n(\omega)$ be the characteristic function of $\overline{X}_n$:

$$M_n(\omega) = \mathrm{E}\left(e^{i\omega(X_1+X_2+\cdots+X_n)/n}\right)$$

$$= \mathrm{E}\left(e^{i\frac{\omega}{n}X_1}\ e^{i\frac{\omega}{n}X_2}\ \cdots\ e^{i\frac{\omega}{n}X_n}\right)$$

$$= \mathrm{E}\left(e^{i\frac{\omega}{n}X_1}\right)\ \mathrm{E}\left(e^{i\frac{\omega}{n}X_2}\right)\ \cdots\ \mathrm{E}\left(e^{i\frac{\omega}{n}X_n}\right)$$

$$= \left(M_X\left(\frac{\omega}{n}\right)\right)^n$$

The characteristic function of $\overline{X}_n$

Expanding $M_X(u)$ as a power series in u:

$$M_X(u) = 1 + imu + \frac{i^2}{2}m_2u^2 + \cdots$$

Therefore

$$M_n(\omega) = \left(M_X\left(\frac{\omega}{n}\right)\right)^n$$

$$= \left(1 + im\left(\frac{\omega}{n}\right) + \frac{i^2}{2}m_2\left(\frac{\omega}{n}\right)^2 + \cdots\right)^n$$

The characteristic function of $\overline{X}_n$

Taking logarithms:

$$\ln\left(M_n(\omega)\right) = n\ \ln\left(1 + im\left(\frac{\omega}{n}\right) + \frac{i^2}{2}m_2\left(\frac{\omega}{n}\right)^2 + \cdots\right)$$

$$= n\ \left(im\left(\frac{\omega}{n}\right) + o\left(\frac{1}{n}\right)\right) = im\omega + \frac{o\left(1/n\right)}{1/n}.$$

Then

$$\ln\left(M_n(\omega)\right) \to im\omega \quad \text{as} \quad n \to \infty$$

The characteristic function of $\overline{X}_n$

Therefore

$$M_n(\omega) \to e^{im\omega} \quad \text{as} \quad n \to \infty$$

This limit function is the characteristic function of a "constant" random variable Y such that

$$\mathrm{P}(Y = m) = 1$$

Continuity theorem for characteristic functions

We say that the sequence $F_1, F_2, \dots, F_n, \dots$ of distribution functions converges to the distribution function F, written

$$F_n \to F$$

if

$$F_n(x) \to F(x)$$

at each point x where F is continuous.

Continuity theorem for characteristic functions

Theorem

Suppose that $F_1, F_2, \dots, F_n, \dots$ is a sequence of distribution functions with corresponding characteristic functions $M_1, M_2, \dots, M_n, \dots$

- *If $F_n \to F$ for some distribution function F with characteristic function M, then $M_n(\omega) \to M(\omega)$ for all ω.*
- *Conversely, if $M(\omega) = \lim_{n\to\infty} M_n(\omega)$ exists and is continuous at $\omega = 0$, then M is the characteristic function of some distribution function F, and $F_n \to F$.*

Convergence of the distribution functions of $\overline{X}_n$

Applying the continuity theorem we can give another interpretation of the convergence to m of the sequence of sample means.

Let F_n be the distribution function of $\overline{X}_n$ and let

$$F(x) = \begin{cases} 0, & x < m \\ 1, & x \geq m \end{cases}$$

be the distribution function of a random variable Y with characteristic function $M(\omega) = e^{im\omega}$. (Notice that $Y = m$ with probability 1.)

Then

$$F_n(x) \to F(x) \quad \text{for all } x \neq m$$

Convergence in distribution

Definition

The sequence $X_1, X_2, \dots, X_n, \dots$ converges in distribution to the random variable X as $n \to \infty$ if $F_{X_n} \to F_X$. That is, if

$$F_{X_n}(x) \to F_X(x) \quad \text{for all } x \in C(F_X),$$

where

$$C(F_X) = \{x \in \mathbb{R} : F(x) \text{ is continuous at } x\}$$

Notation:

$$X_n \xrightarrow{d} X$$

Example:

$$\overline{X}_n \xrightarrow{d} m$$

The central limit theorem

Theorem

Let $X_1, X_2, \dots, X_k, \dots$ be a sequence of independent and identically distributed r.v. with $\mathrm{E}(X_k) = m$ and $\mathrm{Var}(X_k) = \sigma^2$.

Let

$$S_n^\star = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k - m}{\sigma}.$$

Then

$$S_n^\star \xrightarrow{d} N(0,1)$$

That is, for all $x \in \mathbb{R}$,

$$\lim_{n\to\infty} F_{S_n^\star}(x) = F_{N(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

The central limit theorem

- We have

$$\mathrm{E}(S_n^\star) = 0, \quad \mathrm{Var}(S_n^\star) = 1$$

- Notice that $S_n^\star$ is the normalized sample mean $\overline{X}_n$:

$$S_n^\star = \frac{\overline{X}_n - m}{\sigma/\sqrt{n}}$$

$$\overline{X}_n = \frac{\sigma}{\sqrt{n}} S_n^\star + m$$

The central limit theorem

Let M_n be the characteristic function of $S_n^\star$. Therefore

$$\begin{aligned}
M_n(\omega) &= \mathrm{E}\left(e^{i\omega S_n^\star}\right) \\
&= \mathrm{E}\left(e^{i\omega \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k - m}{\sigma}}\right) = \mathrm{E}\left(\prod_{k=1}^{n} e^{i \frac{\omega}{\sqrt{n}} \frac{X_k - m}{\sigma}}\right) \\
&= \prod_{k=1}^{n} \mathrm{E}\left(e^{i \frac{\omega}{\sqrt{n}} \frac{X_k - m}{\sigma}}\right) = \left(M_Z\left(\frac{\omega}{\sqrt{n}}\right)\right)^n
\end{aligned}$$

where M_Z is the characteristic function of

$$Z = \frac{X_1 - m}{\sigma}$$

The central limit theorem

Since

$$m_1 = \mathrm{E}(Z) = 0, \quad m_2 = \mathrm{E}\left(Z^2\right) = 1,$$

the first terms of the series expansion of M_Z are:

$$M_Z(u) = 1 + i m_1 u + \frac{i^2}{2} m_2 u^2 + \cdots = 1 - \frac{1}{2} u^2 + \cdots$$

Therefore

$$M_n(\omega) = \left(M_Z\left(\frac{\omega}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{1}{2}\left(\frac{\omega}{\sqrt{n}}\right)^2 + o\left(\frac{\omega}{\sqrt{n}}\right)^2\right)^n$$

The central limit theorem

Taking logarithms,

$$\ln M_n(\omega) = n \ \ln\left(1 - \frac{1}{2}\left(\frac{\omega}{\sqrt{n}}\right)^2 + o\left(\frac{\omega}{\sqrt{n}}\right)^2\right)$$

$$= n\left(-\frac{1}{2}\left(\frac{\omega}{\sqrt{n}}\right)^2 + o\left(\frac{1}{n}\right)\right) \longrightarrow -\frac{1}{2}\,\omega^2,$$

The central limit theorem

Hence,

$$M_n(\omega) \to e^{-\omega^2/2}$$

This is the characteristic function of a standard normal random variable.

Then, by the continuity theorem,

$$S_n^\star \xrightarrow{d} \mathrm{N}(0,1)$$

WLLN versus CLT

Example: Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent random varibles, uniformly distributed on $[0,1]$.

By the weak law of large numbers we have that

$$\mathrm{P}\left(\left|\overline{X}_n - \frac{1}{2}\right| > \frac{1}{10}\right) \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

WLLN versus CLT

On the other hand, by the CLT:

$$\mathrm{P}\left(\left|\overline{X}_n - \frac{1}{2}\right| > \frac{1}{10}\right) = 1 - \mathrm{P}\left(-\frac{1}{10} \le \overline{X}_n - \frac{1}{2} \le \frac{1}{10}\right)$$

$$= 1 - \mathrm{P}\left(-\frac{\sqrt{12n}}{10} \le \frac{\overline{X}_n - 1/2}{1/\sqrt{12n}} \le \frac{\sqrt{12n}}{10}\right)$$

$$\approx 2\left(1 - F_{N(0,1)}\left(\frac{\sqrt{12n}}{10}\right)\right)$$

- We have taken into account that $(\overline{X}_n - 1/2)/(1/\sqrt{12n})$ converges in distribution to a standard normal. Thus, its distribution function can be approximated by $F_{N(0,1)}$.

De Moivre-Laplace Theorem

If each X_k is a Bernoulli random variable with probability p, then

$$S_n = X_1 + X_2 + \cdots X_n \sim \mathrm{Bin}(n, p)$$

with expected value np and variance npq.

Hence, the CLT implies

Theorem

$$\frac{S_n - np}{\sqrt{npq}} \xrightarrow{d} \mathrm{N}(0,1)$$

De Moivre-Laplace Theorem

For instance,

$$\mathrm{P}(a \leq S_n \leq b) = \mathrm{P}(a - 1 < S_n \leq b)$$

$$= \mathrm{P}\left(\frac{a - np - 1}{\sqrt{npq}} < \frac{S_n - np}{\sqrt{npq}} \leq \frac{b - np}{\sqrt{npq}}\right)$$

$$= F_n\left(\frac{b - np}{\sqrt{npq}}\right) - F_n\left(\frac{a - np - 1}{\sqrt{npq}}\right)$$

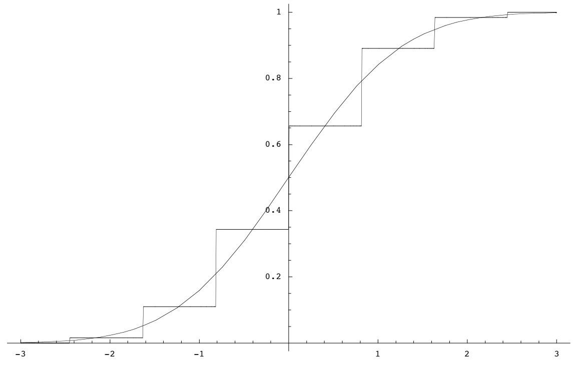
$$\approx F_{N(0,1)}\left(\frac{b - np + 1/2}{\sqrt{npq}}\right) - F_{N(0,1)}\left(\frac{a - np - 1/2}{\sqrt{npq}}\right)$$


```
<< Statistics`DiscreteDistributions`
<< Statistics`ContinuousDistributions`

b6[x_] := CDF[BinomialDistribution[6, 0.5], x]

f6[x_] := Which[x < -3 / Sqrt[6 * 0.25], 0, x < -2 / Sqrt[6 * 0.25], b6[0],
  x < -1 / Sqrt[6 * 0.25], b6[1], x < 0, b6[2], x < 1 / Sqrt[6 * 0.25], b6[3],
  x < 2 / Sqrt[6 * 0.25], b6[4], x < 3 / Sqrt[6 * 0.25], b6[5], x ≥ 3 / Sqrt[6 * 0.25], 1]

Plot[{f6[x], CDF[NormalDistribution[0, 1], x]}, {x, -3, 3}]
```



A local version of the CLT

Theorem

Let $X_1, X_2, \ldots, X_k, \ldots$ be independent identically distributed random variables with zero mean and unit variance, and suppose further that their common characteristic function M_X satisfies

$$\int_{-\infty}^{\infty} |M_X(\omega)|^r \, d\omega < \infty$$

for some integer $r \geq 1$.

Therefore the density f_n of $U_n = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$ exists for $n \geq r$, and

$$f_n(u) \to \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \quad \text{as} \quad n \to \infty, \text{ uniformly in } \mathbb{R}$$

Example: sum of uniform r.v.

Suppose that each X_i is uniform on $[-\sqrt{3}, \sqrt{3}]$. Hence, $\mathrm{E}(X_i) = 0$ and $\mathrm{Var}(X_i) = 1$

Their common characteristic function is

$$
\begin{aligned}
M_X(\omega) &= \int_{-\infty}^{\infty} e^{i\omega x} f_X(x)\, dx \\
&= \frac{1}{2\sqrt{3}} \int_{-\sqrt{3}}^{\sqrt{3}} e^{i\omega x}\, dx = \frac{\sin\left(\sqrt{3}\,\omega\right)}{\sqrt{3}\,\omega}
\end{aligned}
$$

Example: sum of uniform r.v.

We have

$$
\int_{-\infty}^{\infty} |M_X(\omega)|^2\; d\omega = \int_{-\infty}^{\infty} \left(\frac{\sin\left(\sqrt{3}\,\omega\right)}{\sqrt{3}\,\omega}\right)^2\; d\omega = \frac{\pi}{\sqrt{3}} < \infty,
$$

Hence, the sufficient condition of the theorem holds for $r = 2$.

Thus, the density f_n of $U_n = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$ exists for all $n \geq 1$ and

$$
f_n(u) \to \frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}
$$

Example: sum of uniform r.v.

For instance, let

$$
U_3 = \frac{X_1 + X_2 + X_3}{\sqrt{3}}
$$

If

$$
f(s) = f_X(s) * f_X(s) * f_X(s)
$$

then

$$
f_3(u) = \sqrt{3}\, f\left(\sqrt{3}\, s\right)
$$

Example: sum of uniform r.v.

The calculation of f_3 gives

$$
f_3(u) = \begin{cases}
0, & u < -3 \\
(u+3)^2/16, & -3 \leq u < -1 \\
(3-u^2)/8, & -1 \leq u < 1 \\
(3-u)^2/16, & 1 \leq u < 3 \\
0, & u \geq 3
\end{cases}
$$

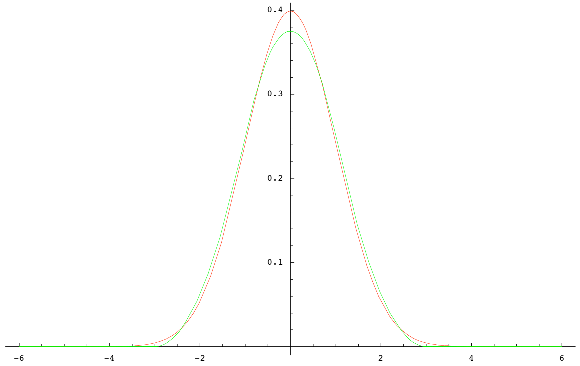
Example: sum of uniform r.v.

Suma de 3 v.a. uniformes en [-Sqrt[3],Sqrt[3]], independents

```
f[t_] := Which[t ≤ -3 Sqrt[3], 0, t ≤ -Sqrt[3],
  (3 Sqrt[3] + t) ^2 / (48 Sqrt[3]), t ≤ Sqrt[3], (9 - t ^2) / (24 Sqrt[3]),
  t ≤ 3 Sqrt[3], (3 Sqrt[3] - t) ^2 / (48 Sqrt[3]), True, 0]

g[t_] := Sqrt[3] f[Sqrt[3] t]

Plot[{1 / Sqrt[2 Pi] Exp[-1 / 2 t ^2], g[t]}, {t, -6, 6},
 PlotRange → All, PlotStyle → {{RGBColor[1, 0, 0]}, {RGBColor[0, 1, 0]}}]
```



Poisson's theorem

Theorem

If $X_n \sim \mathrm{Bin}(n, \lambda/n)$, $\lambda > 0$, *then*

$$X_n \xrightarrow{d} \mathrm{Poiss}(\lambda)$$

Proof:

$$M_{X_n}(\omega) = \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n} e^{i\omega}\right)^n$$
$$= \left(1 + \frac{\lambda(e^{i\omega} - 1)}{n}\right)^n \to e^{\lambda(e^{i\omega}-1)} \quad \text{as} \quad n \to \infty$$

Convergence in distribution: a result

Theorem

Let $X_1, X_2, \dots X_n \dots$ *and* X *be random variables taking nonnegative integer values. A necessary and sufficient condition for* $X_n \xrightarrow{d} X$ *is*

$$\lim_{n\to\infty} \mathrm{P}(X_n = k) = \mathrm{P}(X = k) \quad \textit{for all} \quad k \geq 0$$

For example, by Poisson's theorem and this result, we have that

$$\binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!}$$

Convergence in mean square

Given a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, let us consider the vector space $\mathcal{H}$ whose elements are the random variables X with a finite second order moment, $\mathrm{E}(X^2) < \infty$.

We define an inner product in $\mathcal{H}$ by

$$\langle X, Y \rangle = \mathrm{E}(XY)$$

Convergence in mean square

- The norm induced by this inner product is:

$$\| X \| = \sqrt{\langle X, X \rangle} = \sqrt{\mathrm{E}(X^2)}$$

- Moreover, a distance between random variables of $\mathcal{H}$ can be considered:

$$d(X, Y) = \| X - Y \| = \sqrt{\mathrm{E}\left((X - Y)^2\right)}$$

Convergence in mean square

Definition

The sequence $X_1, X_2, \ldots, X_n, \ldots$ converges in mean square to the random variable X as $n \to \infty$ if

$$d(X_n, X) \to 0 \quad \text{as} \quad n \to \infty,$$

Equivalently,

$$\mathrm{E}\left((X_n - X)^2\right) \to 0 \quad \text{as} \quad n \to \infty$$

Notation:

$$X_n \xrightarrow{2} X$$

Example

The sequence of sample means $\overline{X}_n$ converges in mean square to the expected value m:

$$\mathrm{E}\left(\left(\overline{X}_n - m\right)^2\right) = \mathrm{Var}\left(\overline{X}_n\right) = \frac{\sigma^2}{n} \to 0$$

Convergence in r-mean

More generally,

Definition

The sequence $X_1, X_2, \ldots, X_n, \ldots$ converges in r-mean to the random variable X as $n \to \infty$ if

$$\mathrm{E}\left(|X_n - X|^r\right) \to 0 \quad \text{as} \quad n \to \infty$$

Notation:

$$X_n \xrightarrow{r} X$$

Example

Consider the sequence $X_1, X_2, \ldots, X_n, \ldots$ such that

$$\begin{cases} P(X_n = 0) = 1 - \dfrac{1}{n} \\ P(X_n = -1) = P(X_n = 1) = \dfrac{1}{2n} \end{cases} \qquad n = 1, 2, 3, \ldots$$

Let us prove that, for any $r > 0$,

$$X_n \xrightarrow{r} 0$$

Example

Proof: We have to prove that

$$E\left(|X_n|^r\right) \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty$$

Indeed,

$$E(|X_n|^r) = 0 \cdot P(X_n = 0) + 1 \cdot (P(X_n = -1) + P(X_n = 1))$$

$$= \frac{1}{2n} + \frac{1}{2n} \longrightarrow 0$$

Almost surely convergence

Definition

The sequence $X_1, X_2, \ldots, X_n, \ldots$ converges almost surely to the random variable X as $n \to \infty$ if

$$P\left(\{\omega \in \Omega : X_n(\omega) \to X(\omega) \quad \text{as} \quad n \to \infty\}\right) = 1$$

Notation:

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

We also say that X_n converges to X with probability 1.

Strong laws of large numbers

Theorem

Let $X_1, X_2, \ldots, X_k, \ldots$ be a sequence of independent and identically distributed random variables with finite expectation m and finite variance.

Set

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}, \quad n \geq 1$$

Then,

$$\overline{X}_n \xrightarrow{a.s.} m \quad \textit{as} \quad n \to \infty$$

Strong laws of large numbers

Theorem

Let $X_1, X_2, \ldots, X_k, \ldots$ be a sequence of independent and identically distributed random variables and set

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}, \quad n \geq 1$$

Then

$$\overline{X}_n \xrightarrow{a.s.} \mu \quad \text{as} \quad n \to \infty$$

for some constant μ, if and only if $\mathrm{E}(|X_1|) < \infty$. In this case, $\mu = E(X_1)$.

Uniqueness

Theorem

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of random variables. If the sequence converges

- *almost surely,*
- *in probability,*
- *in r-mean,*
- *or in distribution,*

then the limiting random variable (distribution) is unique

Uniqueness

For example, let us prove the uniqueness of the limit in the case of almost sure convergence.

Suppose that $X_n \xrightarrow{a.s.} X$ and $X_n \xrightarrow{a.s.} Y$ and let

$$N_X = \{\omega : X_n(\omega) \not\to X(\omega) \text{ as } n \to \infty\}$$

$$N_Y = \{\omega : X_n(\omega) \not\to Y(\omega) \text{ as } n \to \infty\}$$

So, we have that

$$\mathrm{P}(N_X) = \mathrm{P}(N_Y) = 0$$

Uniqueness

Let $\omega \in \overline{N_X} \cap \overline{N_Y} = \overline{N_X \cup N_Y}$. Then

$$|X(\omega) - Y(\omega)| \leq |X(\omega) - X_n(\omega)| + |X_n(\omega) - Y(\omega)| \longrightarrow 0$$

So, if $\omega \in \overline{N_X \cup N_Y}$, then $X(\omega) = Y(\omega)$; hence, if $X(\omega) \neq Y(\omega)$, then $\omega \in N_X \cup N_Y$.

Thus

$$\mathrm{P}(X \neq Y) \leq \mathrm{P}(N_X \cup N_Y) \leq \mathrm{P}(N_X) + \mathrm{P}(N_Y) = 0$$

That is,

$$X = Y \quad \text{with probability 1}$$

Relations between the convergence concepts

Theorem

The following implications hold:

- $\left(X_n \xrightarrow{a.s.} X\right) \Longrightarrow \left(X_n \xrightarrow{\mathrm{P}} X\right)$
- $\left(X_n \xrightarrow{r} X\right) \Longrightarrow \left(X_n \xrightarrow{\mathrm{P}} X\right)$ *for any* $r \geq 1$.
- $\left(X_n \xrightarrow{\mathrm{P}} X\right) \Longrightarrow \left(X_n \xrightarrow{d} X\right)$
- *If* $r > s \geq 1$, *then* $\left(X_n \xrightarrow{r} X\right) \Longrightarrow \left(X_n \xrightarrow{s} X\right)$

All implications are strict.

Relations between the convergence concepts

With additional hypothesis some converse implications hold.

Theorem

- *If* c *is a constant, then* $\left(X_n \xrightarrow{d} c\right) \Longrightarrow \left(X_n \xrightarrow{\mathrm{P}} c\right)$
- *If* $X_n \xrightarrow{\mathrm{P}} X$ *and there exists a constant* C *such that* $\mathrm{P}(|X_n| \leq C) = 1$ *for all* n, *then* $X_n \xrightarrow{r} X$ *for all* $r \geq 1$.
- *If* $P_n(\epsilon) = \mathrm{P}\left(|X_n - X| > \epsilon\right)$ *satisfies* $\sum_n P_n(\epsilon) < \infty$ *for all* $\epsilon > 0$, *then* $X_n \xrightarrow{a.s.} X$

Relations between the convergence concepts

For instance, let us consider the proof of the implication

$$\left(X_n \xrightarrow{d} c\right) \Longrightarrow \left(X_n \xrightarrow{\mathrm{P}} c\right)$$

Hence, assume that $X_n \xrightarrow{d} X$ where $X = c$ is a constant r.v. with distribution function

$$F_X(x) = \begin{cases} 0, & x < c \\ 1, & x \geq c \end{cases}$$

Relations between the convergence concepts

If $\epsilon > 0$ is a fixed number, we have that

$$\begin{aligned} P(|X_n - c| > \epsilon) &= 1 - \mathrm{P}(c - \epsilon \leq X_n \leq c + \epsilon) \\ &= 1 - \left(F_{X_n}(c + \epsilon) - F_{X_n}(c - \epsilon) + \mathrm{P}(X_n = c - \epsilon)\right) \\ &\leq 1 - F_{X_n}(c + \epsilon) + F_{X_n}(c - \epsilon) \longrightarrow 0 \end{aligned}$$

because

$$F_{X_n}(c + \epsilon) \to F_X(c + \epsilon) = 1, \quad F_{X_n}(c - \epsilon) \to F_X(c - \epsilon) = 0$$

Therefore

$$X_n \xrightarrow{\mathrm{P}} c$$

Example

Consider the sequence $X_1, X_2, \ldots, X_n, \ldots$ such that

$$P(X_1 = 1) = 1,$$

$$P(X_n = 1) = 1 - \frac{1}{n^2}, \quad P(X_n = n) = \frac{1}{n^2} \quad n \geq 2$$

Let us prove that the sequence converges almost surely to 1.

Example

We have that

$$P_n(\epsilon) = P(|X_n - 1| > \epsilon) = \begin{cases} 0, & n = 1 \\ 1/n^2, & n \geq 2 \end{cases}$$

Therefore

$$\sum_{n=1}^{\infty} P_n(\epsilon) = \sum_{n=2}^{\infty} \frac{1}{n^2} < \infty$$

Hence,

$$X_n \xrightarrow{a.s.} 1$$

Operations with limits

Theorem

- *If* $X_n \xrightarrow{a.s.} X$ *and* $Y_n \xrightarrow{a.s.} Y$*, then* $X_n + Y_n \xrightarrow{a.s.} X + Y$.
- *If* $X_n \xrightarrow{P} X$ *and* $Y_n \xrightarrow{P} Y$*, then* $X_n + Y_n \xrightarrow{P} X + Y$.
- *If* $X_n \xrightarrow{r} X$ *and* $Y_n \xrightarrow{r} Y$ *for some* $r > 0$*, then* $X_n + Y_n \xrightarrow{r} X + Y$.

Cramer's Theorem

Theorem

Suppose that $X_n \xrightarrow{d} X$ *and* $Y_n \xrightarrow{P} a$*, where* a *is a constant. Then*

- $X_n + Y_n \xrightarrow{d} X + a$.
- $X_n - Y_n \xrightarrow{d} X - a$.
- $X_n \cdot Y_n \xrightarrow{d} X \cdot a$.
- $X_n / Y_n \xrightarrow{d} X / a$, *for* $a \neq 0$.

Operations with limits

Theorem

Let $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$. If X_n and Y_n are independent random variables for all n and, moreover, X and Y are independent, then $X_n + Y_n \xrightarrow{d} X + Y$.

Proof:
It suffices to proof that $M_{X_n+Y_n}(\omega) \to M_{X+Y}(\omega)$ as $n \to \infty$ (for $-\infty < \omega < \infty$).

Thus, it suffices to proof that $M_{X_n}(\omega)M_{Y_n}(\omega) \to M_X(\omega)M_Y(\omega)$.

But this is a simple consequence of $M_{X_n}(\omega) \to M_X(\omega)$ and $M_{Y_n}(\omega) \to M_Y(\omega)$.

Continuous functions

Theorem

Let $X_n \xrightarrow{P} a$, where a is a constant. Suppose, further, that g is a continuous function at point a. Then

$$g(X_n) \xrightarrow{P} g(a) \quad \text{as } n \longrightarrow \infty.$$

Continuous functions

Proof:
Given $\epsilon > 0$ there exists $\delta > 0$ such that $|g(x) - g(a)| < \epsilon$ if $|x - a| < \delta$. Hence,

$$\{|g(X_n) - g(a)| \geq \epsilon\} \subset \{|X_n - a| \geq \delta\}$$

and, thus,

$$\mathrm{P}(|g(X_n) - g(a)| \geq \epsilon) \leq \mathrm{P}(|X_n - a| \geq \delta)$$

But $\mathrm{P}(|X_n - a| \geq \delta) \to 0$ because $X_n \xrightarrow{P} a$. Therefore, $\mathrm{P}(|g(X_n) - g(a)| \geq \epsilon) \to 0$ and

$$g(X_n) \xrightarrow{P} g(a) \quad \text{as } n \longrightarrow \infty$$

Delta method

Theorem

Let

(a) *$\{a_n : n \geq 1\}$ be a sequence of real numbers such that $a_n \to \infty$ as $n \to \infty$, and $a_n \neq 0$ for all n,*

(b) *$\{X_n : n \geq 1\}$ be a sequence of random variables an θ be a real number such that $a_n(X_n - \theta) \xrightarrow{d} N(0, \sigma)$,*

(c) *g be a real function with a continuous derivative in an interval that contains theta and such $g'(\theta) \neq 0$.*

Then

$$a_n(g(X_n) - g(\theta)) \xrightarrow{d} N(0, |g'(\theta)|\sigma).$$

Shorokhod's representation theorem

Theorem

Let $X_1, X_2, \dots, X_k, \dots$ and X be random variables such that $X_n \xrightarrow{d} X$ as $n \to \infty$.

Then there exists a probability space $(\Omega', \mathcal{F}', \mathrm{P}')$ and random variables $Y_1, Y_2, \dots, Y_k, \dots$ and Y, which map Ω' into $\mathbb{R}$, such that:

- *$Y_1, Y_2, \dots, Y_k, \dots$ and Y have the same distribution functions that $X_1, X_2, \dots, X_k, \dots$ and X, respectively.*
- *$Y_n \xrightarrow{a.s} Y$.*

Borel-Cantelli lemmas

Given a sequence of events $A_1, A_2, \dots, A_n, \dots$, let

$$B_n = \bigcup_{k=n}^{\infty} A_k, \quad n \geq 1$$

Notice that

$$B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots$$

is a decreasing sequence of events.

Borel-Cantelli lemmas

The limit of the sequence $B_1, B_2, \dots, B_n, \dots$ is

$$A^\star = \lim_{n\to\infty} B_n = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

The event $A^\star$ is called the limit superior of the sequence $A_1, A_2, \dots, A_n, \dots$, and it is denoted by $A^\star = \limsup_n A_n$.

- Notice that $\omega \in A^\star$ if and only if ω belongs to infinitely many of the A_n.
- That is, $A^\star$ is the event "infinitely many of the A_n occur".

Borel-Cantelli lemmas

Theorem (Borel-Cantelli lemmas)

Let $A_1, A_2, \dots, A_n, \dots$ be a sequence of events and $A^\star$ its limit superior. Then:

- $\mathrm{P}(A^\star) = 0$ *if* $\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty$,
- $\mathrm{P}(A^\star) = 1$ *if* $\sum_{n=1}^{\infty} \mathrm{P}(A_n) = \infty$ *and the events* $A_1, A_2, \dots$ *are independent.*

Borel-Cantelli lemmas

For example, the proof of the first Borel-Cantelli lemma is:

$$P(A^\star) = \mathrm{P}\left(\lim_{n\to\infty} B_n\right) = \lim_{n\to\infty} \mathrm{P}(B_n)$$
$$= \lim_{n\to\infty} \mathrm{P}\left(\bigcup_{k=n}^{\infty} A_k\right) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathrm{P}(A_k) = 0$$

because $\sum_n \mathrm{P}(A_n)$ converges by hypothesis.

Zero-one law

Corollary (zero-one law)

Let $A_1, A_2, \dots, A_n, \dots$ be a sequence of independent events and let $A^\star$ be its limit superior.

Then either $\mathrm{P}(A^\star) = 0$ or $\mathrm{P}(A^\star) = 1$ according as $\sum_{n=1}^{\infty} \mathrm{P}(A_n)$ converges or diverges respectively.

Example

Let $X_1, X_2, \dots, X_n, \dots$ be a sequence of random variables such that

$$\mathrm{P}(X_n = 0) = \frac{1}{n^2}, \quad \mathrm{P}(X_n = 1) = 1 - \frac{1}{n^2} \quad n \ge 1$$

Let us consider the events $A_n = \{X_n = 0\}$, $n \ge 1$.

- Since $\sum_n \mathrm{P}(A_n) < \infty$, the first Borel-Cantelli lemma implies that $\mathrm{P}(A^\star) = 0$. Thus, there is a 0 probability that $\{X_n = 0\}$ happens infinitely often.
- Therefore $\mathrm{P}(X_n = 1$ for all n suficiently large$) = 1$.

  Thus we have proved:

$$\lim_{n\to\infty} X_n = 1 \quad \text{with probability } 1$$

Example

Now, let $X_1, X_2, \dots, X_n, \dots$ be independent random variables such that

$$\mathrm{P}(X_n = 0) = \frac{1}{n}, \quad \mathrm{P}(X_n = 1) = 1 - \frac{1}{n} \quad n \ge 1$$

and let $A_n = \{X_n = 0\}$, $n \ge 1$.

- Since $\sum_n \mathrm{P}(A_n) = \infty$, the second Borel-Cantelli lemma implies that, with probability 1, infinitely many of the events $A_n = \{X_n = 0\}$ occur.
- Analogously, $\sum_n \mathrm{P}(\overline{A}_n) = \infty$. Thus, the probability that infinitely many of the events $\overline{A}_n = \{X_n = 1\}$ occur is also 1.
- Hence, with probability 1, $\lim_{n\to\infty} X_n$ does not exist.