STK4011 and STK9011
Autumn 2016

Convergence of random variables

Covers (most of) the following material from chapters 2, 3 and 5:

- Section 2.3: page 66
- Section 3.6.1: page 122
- Sections 5.5.1-3: pages 232-240

Ørnulf Borgan
Department of Mathematics
University of Oslo

Convergence in probability

We will look at different concepts of convergence for a sequence $X_1, X_2, \ldots$ of random variables

Definition 5.5.1
A sequence $X_1, X_2, \ldots$ of random variables converges in probability to a random variable X if for every $\varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

or equivalently

$$\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

Chebychev's inequality will be useful:

Theorem 3.6.1 (Chebychev's inequality)
Let X be a random variable and let $g(x)$ be a non-negative function. Then, for every $r > 0$,

$$P\big(g(X) \geq r\big) \leq \frac{\mathrm{E}g(X)}{r}$$

The following is a much used version of the inequality:

$$P(|X - \mathrm{E}X| \geq \varepsilon) \leq \frac{\mathrm{Var}X}{\varepsilon^2}$$

Theorem 5.5.2 (weak law of large numbers)
Let $X_1, X_2, \ldots$ be iid random variables with $\mathrm{E}X_i = \mu$ and $\mathrm{Var}X_i = \sigma^2$

Define

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Then for every $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| < \varepsilon) = 1$$

That is, $\bar{X}_n$ converges in probability to μ

We say that $\bar{X}_n$ is a consistent estimator of μ

Example 5.5.3 (consistency of S^2)

Let $X_1, X_2, ...$ be iid random variables with $EX_i = \mu$ and $\text{Var}X_i = \sigma^2$. We assume that the 4th moment of the X_i's is finite.

Define

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

We have:

- $ES_n^2 = \sigma^2$ (page 214)
- $\text{Var}S_n^2 \to 0$ (exercise 5.8.b)

Then for every $\varepsilon > 0$

$$P(|S_n^2 - \sigma^2| \geq \varepsilon) \leq \frac{\text{Var}S_n^2}{\varepsilon^2} \to 0 \quad \text{as} \quad n \to \infty$$

Theorem 5.5.4 (modified)

Assume that $X_1, X_2, ...$ converges in probability to a constant a, and let $h(x)$ be a function that is continuous at $x = a$. Then $h(X_1), h(X_2), ...$ converges in probability to $h(a)$.

Proof

Let $\varepsilon > 0$. Since $h(x)$ is continuous at $x = a$, there exist a $\delta > 0$ such that

$$|x-a| < \delta \quad \Rightarrow \quad |h(x) - h(a)| < \varepsilon$$

Thus

$$|h(x) - h(a)| \geq \varepsilon \quad \Rightarrow \quad |x-a| \geq \delta$$

It follows that (as $n \to \infty$)

$$P\left(|h(X_n) - h(a)| \geq \varepsilon\right) \leq P\left(|X_n - a| \geq \delta\right) \to 0$$

Example 5.5.5 (consistency of S)

Let $X_1, X_2, ...$ be iid random variables with $EX_i = \mu$ and $\text{Var}X_i = \sigma^2$. We assume that the 4th moment of the X_i's is finite.

Then

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$$

is a consistent estimator of σ^2

It follows that

$$S_n = \sqrt{S_n^2}$$

is a consistent estimator of $\sqrt{\sigma^2} = \sigma$

Almost sure convergence

Definition 5.5.6

A sequence $X_1, X_2, ...$ of random variables converges almost surely (or with probability one) to a random variable X if

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

Theorem 5.5.9 (strong law of large numbers)

Let $X_1, X_2, ...$ be iid random variables with $EX_i = \mu$. Then

$$P(\lim_{n\to\infty} \bar{X}_n = \mu) = 1$$

That is, $\bar{X}_n$ converges almost surely to μ

Convergence in distribution

Definition 5.5.10
A sequence $X_1, X_2, \ldots$ of random variables with cdfs $F_{X_1}(x), F_{X_2}(x), \ldots$ converges in distribution to a random variable X if

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

at all points x where the cdf $F_X(x)$ of X is continuous

Example 5.5.11 (maximum of uniforms)
Let $X_1, X_2, \ldots$ be iid uniform(0,1)

Consider $X_{(n)} = \max_{1\le i\le n} X_i$

Then:

- $X_{(n)}$ converges to 1 in probability
- $n\left(1 - X_{(n)}\right)$ converges in distribution to X, where $X \sim \text{exponential}(1)$

Theorem 5.5.13
A sequence $X_1, X_2, \ldots$ of random variables converges in probability to a constant a if and only if the sequence converge in distribution to the one-point distribution in a

That is, the statement

$$P(|X_n - a| \ge \varepsilon) \to 0 \quad \text{for every } \varepsilon > 0$$

is equivalent to

$$P(X_n \le x) \to \begin{cases} 0 & \text{if } x < a \\ 1 & \text{if } x > a \end{cases}$$

Theorem 2.3.12
Suppose that $X_1, X_2, \ldots$ is a sequence of random variables and that X_i has mgf $M_{X_i}(t)$ and cdf $F_{X_i}(x),\ i = 1, 2, \ldots$

Furthermore, suppose that

$$\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$$

for all t in a neighbourhood of 0, where $M_X(t)$ is the mgf of a random variable X with cdf $F_X(x)$

Then

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

for all x where $F_X(x)$ is continuous

That is, $X_1, X_2, \ldots$ converges in distribution to X

Theorem 5.5.15 (Central limit theorem)

Let $X_1, X_2, \ldots\ldots$ be iid random variables with $EX_i = \mu$ and $\mathrm{Var}X_i = \sigma^2$

Let $G_n(x)$ be the cdf of $\sqrt{n}(\bar{X}_n - \mu)/\sigma$

Then, for any x ($-\infty < x < \infty$),

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

That is, $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ has a limiting standard normal distribution

Definition 5.5.17 (Slutsky's theorem)

If $X_n \to X$ in distribution and $Y_n \to a$ in probability, where a is a constant, then

a) $Y_n X_n \to aX$ in distribution

b) $X_n + Y_n \to X + a$ in distribution

Example 5.5.18

Let $X_1, X_2, \ldots\ldots$ be iid random variables with $EX_i = \mu$ and $\mathrm{Var}X_i = \sigma^2$. Then

$\sqrt{n}(\bar{X}_n - \mu)/\sigma \to n(0,1)$ in distribution

$\sigma / S_n \to 1$ in probability

It follows that

$\sqrt{n}(\bar{X}_n - \mu)/S_n \to n(0,1)$ in distribution

Proof of Slutsky's theorem

We will first prove b)

It is sufficient to consider the case where $a = 0$

We have to prove that

$$\lim_{n\to\infty} F_{X_n+Y_n}(x) = F_X(x) \qquad (*)$$

when x is a continuity point of $F_X(x)$

For all $\varepsilon > 0$ we have

$$F_{X_n+Y_n}(x) = P(X_n + Y_n \le x) = P(X_n \le x - Y_n)$$

$$= P\big((X_n \le x - Y_n) \cap (|Y_n| \ge \varepsilon)\big) + P\big((X_n \le x - Y_n) \cap (|Y_n| < \varepsilon)\big)$$

$$= Q_{n1} + Q_{n2}$$

Now $0 \le Q_{n1} \le P\big(|Y_n| \ge \varepsilon\big) \to 0$, so $Q_{n1} \to 0$

Furthermore

$$Q_{n2} \le P\big((X_n \le x + \varepsilon) \cap (|Y_n| < \varepsilon)\big) \le P\big(X_n \le x + \varepsilon\big) = F_{X_n}(x + \varepsilon)$$

and

$$Q_{n2} \ge P\big((X_n \le x - \varepsilon) \cap (|Y_n| < \varepsilon)\big)$$
$$= 1 - P\big((X_n \le x - \varepsilon)^c \cup (|Y_n| < \varepsilon)^c\big)$$
$$\ge 1 - P(X_n > x - \varepsilon) - P(|Y_n| \ge \varepsilon)$$
$$= F_{X_n}(x - \varepsilon) - P(|Y_n| \ge \varepsilon)$$

Thus

$$F_{X_n}(x - \varepsilon) - P(|Y_n| \ge \varepsilon) \le Q_{n2} \le F_{X_n}(x + \varepsilon)$$

We may assume that $x \pm \varepsilon$ are continuity points of $F_X(x)$ (since one may show that the number of discontinuity points is countable)

Then we obtain

$$F_X(x-\varepsilon) \le \liminf Q_{n2} \le \limsup Q_{n2} \le F_X(x+\varepsilon)$$

Since we may choose ε arbitrary small and x is a continuity point of $F_X(x)$, this shows that $Q_{n2} \to F_X(x)$, and hence (*) is proved

We will then prove a)

We first consider the case when $a = 0$

By Theorem 5.5.13 we then have to prove that $Y_n X_n \to 0$ in probability

Note first that $P(|X| \ge K)$ can be made arbitrarily small by choosing K large enough

Now we have for all $\delta, \varepsilon > 0$ that

$$P(|Y_n X_n| \ge \varepsilon) = P\big((|Y_n X_n| \ge \varepsilon) \cap (|Y_n| \ge \delta)\big)$$
$$+ P\big((|Y_n X_n| \ge \varepsilon) \cap (|Y_n| < \delta)\big)$$
$$\le P(|Y_n| \ge \delta) + P(|X_n| \ge \varepsilon/\delta)$$

We may assume that $\pm\varepsilon/\delta$ are continuity points of $F_X(x)$

Then we obtain

$$0 \le \liminf P(|Y_n X_n| \ge \varepsilon) \le \limsup P(|Y_n X_n| \ge \varepsilon) \le P(|X| \ge \varepsilon/\delta)$$

Here the right-hand side can be made arbitrary small by choosing d small enough, and it follows that $\lim P(|Y_n X_n| \ge \varepsilon) = 0$, i.e. $Y_n X_n \to 0$ in probability

If` $a \ne 0$ we may write $X_n Y_n = aX_n + X_n(Y_n - a)$

By what has just been proved, we have that $X_n(Y_n - a) \to 0$ in probability

By result b) of the theorem, it then only remains to prove that $aX_n \to aX$ in distribution

If $a > 0$ we have when x/a is a continuity point of $F_X(x)$

$$P(aX_n \le x) = P(X_n \le x/a) \to P(X \le x/a) = P(aX \le x)$$

and it follows that $aX_n \to aX$ in distribution

Similarly, if $a < 0$ we have when x/a is a continuity point of $F_X(x)$

$$P(aX_n \le x) = 1 - P(aX_n > x) = 1 - P(X_n < x/a)$$
$$\to 1 - P(X < x/a) = 1 - P(aX > x) = P(aX \le x)$$

and it follows that $aX_n \to aX$ in distribution