Stefan Grosskinsky
www.statslab.cam.ac.uk/~stefan/teaching/probmeas.html
Part IID, M 2005

Probability and Measure

Hand-out 2

Convergence of random variables

Let $X, X_n : \Omega \to \mathbb{R}$ be random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.
There are several concepts of convergence of random variables, which are summarised in the following:

(i) $X_n \to X$ *everywhere* or *pointwise* if $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega$ as $n \to \infty$.

(ii) $X_n \xrightarrow{a.s.} X$ *almost surely* (*a.s.*) if $\mathbb{P}(X_n \not\to X) = 0$.

(iii) $X_n \xrightarrow{P} X$ *in probability* if $\forall\, \epsilon > 0 \; : \; \mathbb{P}(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$.

(iv) $X_n \xrightarrow{L^p} X$ *in* L^p for $p \in [1, \infty]$, if $\|X_n - X\|_p \to 0$ as $n \to \infty$.

(v) $X_n \xrightarrow{D} X$ in *distribution* if $\mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x)$ as $n \to \infty$, for all continuity points of $\mathbb{P}(X \leq x)$. This is also called *weak convergence* or *convergence in law*.

The following implications hold $(q \geq p \geq 1)$:

$$\begin{array}{ccccccc}
X_n \to X & \Rightarrow & X_n \xrightarrow{a.s.} X & \searrow & & & \\
 & & & & X_n \xrightarrow{P} X & \Rightarrow & X_n \xrightarrow{D} X \\
X_n \xrightarrow{L^q} X & \Rightarrow & X_n \xrightarrow{L^p} X & \nearrow & & &
\end{array}$$

Proofs are given in Theorem 3.10, Corollary 5.3 and problem sheet question 3.2.
The last implication is proved below.

Proof. of $X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{D} X$:
Suppose $X_n \xrightarrow{P} X$ and write $F_n(x) = \mathbb{P}(X_n \leq x)\,,\ F(x) = \mathbb{P}(X \leq x)$ for the distribution functions.
If $\epsilon > 0$, $F_n(x) = \mathbb{P}(X_n \leq x,\ X \leq x + \epsilon) + \mathbb{P}(X_n \leq x,\ X > x + \epsilon) \leq F(x + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon)$.
Similarly, $F(x - \epsilon) = \mathbb{P}(X \leq x - \epsilon,\ X_n \leq x) + \mathbb{P}(X \leq x - \epsilon,\ X_n > x) \leq F_n(x) + \mathbb{P}(|X_n - X| > \epsilon)$.
Thus $F(x - \epsilon) - \mathbb{P}(|X_n - X| > \epsilon) \leq F_n(x) \leq F(x + \epsilon) + \mathbb{P}(|X_n - X| > \epsilon)$, and as $n \to \infty$ we obtain

$$F(x - \epsilon) \leq \liminf_n F_n(x) \leq \limsup_n F_n(x) \leq F(x + \epsilon) \quad \text{for all } \epsilon > 0\,.$$

If F is continuous at x, $F(x - \epsilon) \nearrow F(x)$ and $F(x + \epsilon) \searrow F(x)$ as $\epsilon \to 0$, proving the result. $\square$

Connection between Lebesgue and Riemann integration

Definition. $f : [a, b] \to \mathbb{R}$ is *Riemann integrable* (R-integrable) with integral $R \in \mathbb{R}$, if

$$\forall\, \epsilon > 0 \;\; \exists\, \delta > 0 \;\; \forall\, x_j \in I_j \; : \; \Big| R - \sum_{j=1}^{n} f(x_i)\, |I_j| \Big| < \epsilon \,, \tag{1}$$

for some finite partition $\{I_1, \ldots, I_n\}$ of $[a, b]$ into subintervals of lengths $|I_j| < \delta$.

This corresponds to an approximation of f by step functions, a special case of simple functions which are used defining the Lebesgue integral.

Theorem. Lebesgue's integrability criterium
$f : [a, b] \to \mathbb{R}$ is R-integrable if and only if f is bounded on $[a, b]$ and continuous almost everywhere, i.e. the set of points in $[a, b]$ where f is not continuous has L-measure 0.

Corollary. *Suppose $f : [a, b] \to \mathbb{R}$ is R-integrable. Then it is also Lebesgue integrable (L-integrable) and the values of both integrals coincide.*

Proof. For a partition $\{I_1, \ldots, I_n\}$ define the simple functions

$$\overline{g}_n = \sum_{j=1}^{n} \sup\{f(x) : x \in I_j\}\, \mathbb{1}_{I_j} \,, \quad \underline{g}_n = \sum_{j=1}^{n} \inf\{f(x) : x \in I_j\}\, \mathbb{1}_{I_j} \,.$$

Thus $\underline{g}_n \le f \le \overline{g}_n$ and if f is continuous *a.e.*, $\underline{g}_n, \overline{g}_n \to f$ *a.e.* as $n \to \infty$ and $|I_j| \to 0$.
Since f is bounded, it follows by dominated convergence for the L-integrals

$$\int_a^b \underline{g}_n(x)\, dx \to \int_a^b f(x)\, dx \,, \quad \int_a^b \overline{g}_n(x)\, dx \to \int_a^b f(x)\, dx \,,$$

so the sum in (1) converges and $R = \int_a^b f(x)\, dx$.
On the other hand, if the sum in (1) converges, then $\int_a^b \big|\overline{g}_n(x) - \underline{g}_n(x)\big|\, dx \to 0$ and thus for the limit functions $\underline{g} = \overline{g}$ *a.e.*, and f is continuous *a.e.* since $\underline{g} \le f \le \overline{g}$.
If f was not bounded, one could choose x_j in (1) such that the sum does not converge. □

On the other hand, not every L-integrable function is also R-integrable. The standard example is $f = \mathbb{1}_{[0,1]\cap\mathbb{Q}}$, which can be made R-integrable by changing it on a set of L-measure 0.
This might suggest that for every L-integrable f there exists an R-integrable g with $f = g$ *a.e.* .

This is not true as demonstrated by the following example:
Let $\{r_1, r_2 \ldots\}$ be an enumeration of the rationals in $(0, 1)$. For small $\epsilon > 0$ and each $n \in \mathbb{N}$ choose an open interval $I_n \subseteq (0, 1)$ with $r_n \in I_n$ and L-measure $\lambda(I_n) < \epsilon\, 2^{-n}$. Put $A = \bigcup_n I_n$. Then A is dense in $(0, 1)$ with $0 < \lambda(A) < \epsilon$ and thus for any non-degenerate subinterval I of $(0, 1)$, $\lambda(A \cap I) > 0$.
Take $f = \mathbb{1}_A$ and suppose that $f = g$ a.e.. Let $\{I_j\}$ be some decomposition of $(0, 1)$ into subintervals. Since for each j, $\lambda(I_j \cap A \cap \{f = g\}) = \lambda(I_j \cap A) > 0$, $g(x_j) = f(x_j) = 1$ for some $x_j \in I_j \cap A$, and thus

$$\sum_{j=1}^{n} g(x_j)\, \lambda(I_j) = 1 > \lambda(A) \,. \tag{2}$$

If g were R-integrable, its integral would have to coincide with the L-integral $\int_0^1 f\, d\lambda = \lambda(A)$, which is in contradiction to (2).