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 $(\Omega, \mathcal{A}, \mathbb{P})$ with distributions $\mu, \mu_n$.
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* or in *law* if $\mathbb{P}(X_n \le x) \to \mathbb{P}(X \le x)$ as $n \to \infty$, for all continuity points of $\mathbb{P}(X \le x)$. Since equivalent to (vi), this is often also called *weak convergence*.

(vi) $\mu_n \Rightarrow \mu$ *weakly* if $\displaystyle\int_{\mathbb{R}} f\, d\mu_n \to \int_{\mathbb{R}} f\, d\mu$ for all $f \in C_b(\mathbb{R}, \mathbb{R})$ .

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

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

Proofs are given in Theorem 2.10, Proposition 2.11 (see below), Theorem 3.9 (see next page), Corollary 5.3 and example sheet question 3.2.

**Proof.** of Proposition 2.11: $X_n \xrightarrow{P} X \quad\Rightarrow\quad X_n \xrightarrow{D} X$
Suppose $X_n \xrightarrow{P} X$ and write $F_n(x) = \mathbb{P}(X_n \le x)$ , $F(x) = \mathbb{P}(X \le x)$ for the distr. fcts.
If $\epsilon > 0$, $F_n(x) = \mathbb{P}\big(X_n \le x,\ X \le x+\epsilon\big) + \mathbb{P}\big(X_n \le x,\ X > x+\epsilon\big) \le F(x+\epsilon) + \mathbb{P}(|X_n - X| > \epsilon)$ .
Similarly, $F(x-\epsilon) = \mathbb{P}\big(X \le x-\epsilon,\ X_n \le x\big) + \mathbb{P}\big(X \le x-\epsilon,\ X_n > x\big) \le F_n(x) + \mathbb{P}(|X_n - X| > \epsilon)$.
Thus $F(x-\epsilon) - \mathbb{P}(|X_n - X| > \epsilon) \le F_n(x) \le F(x+\epsilon) + \mathbb{P}(|X_n - X| > \epsilon)$ , and as $n \to \infty$

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

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. □

**Proof.** of Theorem 3.9: $X_n \xrightarrow{D} X \quad \Leftrightarrow \quad \mu_n \Rightarrow \mu$

Suppose $X_n \xrightarrow{D} X$. Then by the Skorohod theorem 2.12 there exist $Y \sim X$ and $Y_n \sim X_n$ on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that, $\quad f(Y_n) \to f(Y)\ a.e. \quad$ since $f \in C_b(\mathbb{R}, \mathbb{R})$. Thus

$$\int_{\mathbb{R}} f\, d\mu_n = \int_\Omega f(Y_n)\, d\mathbb{P} \to \int_\Omega f(Y)\, d\mathbb{P} = \int_{\mathbb{R}} f\, d\mu \quad \text{and} \quad \mu_n \Rightarrow \mu \quad \text{by bounded convergence .}$$

Suppose $\mu_n \to \mu$ and let $y$ be a continuity point of $F_X$.

For $\delta > 0$, approximate $\mathbb{1}_{(-\infty,y]}$ by $\quad f_\delta(x) = \begin{cases} \mathbb{1}_{(-\infty,y]}(x) & , \ x \notin (y, y+\delta) \\ 1 + (y-x)/\delta & , \ x \in (y, y+\delta) \end{cases} \quad$ such that

$$\Big| \int_{\mathbb{R}} (\mathbb{1}_{(-\infty,y]} - f_\delta)\, d\mu \Big| \le \Big| \int_{\mathbb{R}} g_\delta\, d\mu \Big| \quad \text{where} \quad g_\delta(x) = \begin{cases} 1+(x-y)/\delta & , \ x \notin (y-\delta, y) \\ 1+(y-x)/\delta & , \ x \in [y, y+\delta) \\ 0 & , \ \text{otherwise} \end{cases} .$$

The same inequality holds for $\mu_n$ for all $n \in \mathbb{N}$. Then as $n \to \infty$

$$\big|F_{X_n}(y) - F_X(y)\big| = \Big| \int_{\mathbb{R}} \mathbb{1}_{(-\infty,y]}\, d\mu_n - \int_{\mathbb{R}} \mathbb{1}_{(-\infty,y]}\, d\mu \Big| \le$$

$$\le \Big| \int_{\mathbb{R}} g_\delta\, d\mu_n \Big| + \Big| \int_{\mathbb{R}} g_\delta\, d\mu \Big| + \Big| \int_{\mathbb{R}} f_\delta\, d\mu_n - \int_{\mathbb{R}} f_\delta\, d\mu \Big| \to 2 \Big| \int_{\mathbb{R}} g_\delta\, d\mu \Big| ,$$

since $f_\delta, g_\delta \in C_b(\mathbb{R}, \mathbb{R})$. Now, $\quad \big| \int_{\mathbb{R}} g_\delta\, d\mu \big| \le \mu\big((y-\delta, y+\delta)\big) \to 0$ as $\delta \to 0$ , $\quad$ since $\mu\big(\{y\}\big) = 0$, so $X_n \xrightarrow{D} X$. □

**Skorohod representation theorem**

*For all probability distribution functions $F_1, F_2, \ldots : \mathbb{R} \to [0,1]$ there exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and random variables $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ such that $X_n$ has distribution function $F_n$.*

*(a) The $X_n$ can be chosen to be independent.*

*(b) If $F_n \to F$ for all continuity points of the probability distribution function $F$, then the $X_n$ can also be chosen such that $X_n \to X\ a.s.$ with $X : \Omega \to \mathbb{R}$ having distribution function $F$.*

**Proof.** Consider the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ where $\Omega = (0,1]$, $\mathcal{A} = \mathcal{B}\big((0,1]\big)$ and $\mathbb{P}$ is the restriction of Lebesgue measure to $\mathcal{A}$. For each $n \in \mathbb{N}$ define $G_n : (0,1] \to \mathbb{R}$ , $\quad G_n(\omega) = \inf\big\{x \ : \ \omega \le F_n(x)\big\}$.

(b) In problem 2.4 it is shown that $X_n = G_n$ are random variables with distribution functions $F_n$ and that $X_n \to X\ a.s.$ under the assumption in (b), where $X(\omega) = G(\omega)$ is defined analogously.

(a) Each $\omega \in \Omega$ has a unique binary expansion $\omega = 0.\omega_1\omega_2\omega_3\ldots$, where we forbid infinite sequences of 0's. The **Rademacher functions** $R_n : \Omega \to \{0,1\}$ are defined as $R_n(\omega) = \omega_n$. Note that

$$R_1 = \mathbb{1}_{(\frac{1}{2},1]} , \quad R_2 = \mathbb{1}_{(\frac{1}{4},\frac{1}{2}]} + \mathbb{1}_{(\frac{3}{4},1]} , \quad R_3 = \mathbb{1}_{(\frac{1}{8},\frac{1}{4}]} + \mathbb{1}_{(\frac{3}{8},\frac{1}{2}]} + \mathbb{1}_{(\frac{5}{8},\frac{3}{4}]} + \mathbb{1}_{(\frac{7}{8},1]} , \quad \ldots$$

thus in general $\quad R_n = \mathbb{1}_{A_n} \quad$ where $\quad A_n = \bigcup_{k=1}^{2^{n-1}} I_{n,k} \quad$ and $\quad I_{n,k} = \Big(\frac{2k-1}{2^n}, \frac{2k}{2^n}\Big]$ .

With problem 1.11 the $R_n$ are independent if and only if the $A_n$ are. To see this, take $n_1 < \ldots < n_L$ for some $L \in \mathbb{N}$ and we see that $A_{n_1}, \ldots, A_{n_L}$ are independent by induction, using that

$$\mathbb{P}(I_{n_l,k} \cap A_{n_{l+1}}) = \tfrac{1}{2}\mathbb{P}(I_{n_l,k}) = \mathbb{P}(I_{n_l,k})\, \mathbb{P}(A_{n_{l+1}}) \quad \text{for all } k = 1, \ldots, 2^{n_l - 1} ,$$

and thus $\qquad \mathbb{P}(A_{n_1} \cap \ldots \cap A_{n_l} \cap A_{n_{l+1}}) = \mathbb{P}(A_{n_1} \cap \ldots \cap A_{n_l})\, \mathbb{P}(A_{n_{l+1}})$ .

Now choose a bijection $\quad m : \mathbb{N}^2 \to \mathbb{N} \quad$ and set $\quad Y_{k,n} = R_{m(k,n)} \quad$ and $\quad Y_n = \sum_{k=1}^{\infty} 2^{-k} Y_{k,n}$ .

Then $Y_1, Y_2, \ldots$ are independent and $\quad \mathbb{P}\big(i\, 2^{-k} < Y_n \le (i+1)\, 2^{-k}\big) = 2^{-k} \quad$ for all $n, k, i$ .

Thus $\mathbb{P}(Y_n \le x) = x$ for all $x \in (0,1]$. So $X_n = G_n(Y_n)$ are independent random variables with distribution $F_n$, which can be shown analogous to (b). □