Advanced probability I
Lecture 9

Introduction to weak convergence
Weak convergence and convergence of measures

Introduction to weak convergence

Let (X_n) be a sequence of random variables, and let X be some other variable. In Chapter 1, we introduced weak convergence as follows: X_n converges weakly to X if $\lim_{n\to\infty} Ef(X_n) = Ef(X)$ for all continuous, bounded $f : \mathbb{R} \to \mathbb{R}$.

In fact, weak convergence is more naturally discussed in terms of **sequences of probability measures** (μ_n) instead of **sequences of random variables** (X_n). This is because weak convergence is about the **distribution** of X_n moving closer to the **distribution** of X, and not any sort of "pointwise"-like convergence of X_n to X.

Introduction to weak convergence

Let $\mathbb{M}(\mathcal{B})$ denote the set of probability measures on $(\mathbb{R}, \mathcal{B})$. **Weak convergence of probability measures** is a way of making sense of convergence in the space $\mathbb{M}(\mathcal{B})$.

Problem: $\mathbb{M}(\mathcal{B})$ is a very high-dimensional space: any element $\mu \in \mathbb{M}(\mathcal{B})$ contains much information. Therefore, measuring distance between two probability measures $\mu, \nu \in \mathbb{M}(\mathcal{B})$ is a complicated task.

Introduction to weak convergence

Idea: Just as the uniform norm is a way of measuring distance on, say, $C([a, b])$, the space of continuous functions $f : [a, b] \to \mathbb{R}$, we might define

$$\|\mu\|_\infty = \sup_{B \in \mathcal{B}} \mu(B),$$

a sort of "uniform norm" on $\mathbb{M}(\mathcal{B})$. Does this lead to interesting results?

Example. Let (x_n) be a sequence in $\mathbb{R}$ converging to x with $x_n \neq x$ for all $n \geq 1$, let $\mu_n = \varepsilon_{x_n}$ for $n \geq 1$ and $\mu = \varepsilon_x$, where ε_x is the Dirac measure at x. Then μ_n does **not** converge to μ in $\| \cdot \|_\infty$.

It appears that $\| \cdot \|_\infty$ is too "strong" a norm to obtain reasonable results.

Introduction to weak convergence

An analogy. Consider convergence in $\mathbb{R}^d$. Let $\hat{X}_i : \mathbb{R}^d \to \mathbb{R}$ denote the i'th coordinate projection. For a sequence (x_n) in $\mathbb{R}^d$, we have:

$$\lim_{n\to\infty} x_n = x \text{ if and only if}$$
$$\lim_{n\to\infty} \hat{X}_i(x_n) = \hat{X}_i(x) \text{ for } i \leq d.$$

Thus, convergence of x_n in $\mathbb{R}^d$ can be seen as convergence of **functions** of x_n, where the set of functions $\{\hat{X}_1, \ldots, \hat{X}_d\}$ considered suffices to "describe" all of x_n.

Can we define a similar convergence concept on $\mathbb{M}(\mathcal{B})$ using an appropriate family of functions $\mathbb{M}(\mathcal{B}) \mapsto \mathbb{R}$?

Introduction to weak convergence

Let $C_b(\mathbb{R})$ denote the set of bounded, continuous functions from $\mathbb{R}$ to $\mathbb{R}$. For each $f \in C_b(\mathbb{R})$, define $\pi_f : \mathbb{M}(\mathcal{B}) \to \mathbb{R}$ by

$$\pi_f(\mu) = \int f \, \mathrm{d}\mu.$$

For any $\mu \in \mathbb{M}(\mathcal{B})$, the family $(\pi_f(\mu))_{f \in C_b(\mathbb{R})}$ would appear to "describe" much of the measure μ. We might then say that for a sequence (μ_n) in $\mathbb{M}(\mathcal{B})$ and some other $\mu \in \mathbb{M}(\mathcal{B})$, μ_n converges to μ if

$$\lim_{n \to \infty} \pi_f(\mu_n) = \pi_f(\mu) \text{ for all } f \in C_b(\mathbb{R}).$$

This turns out to be a good idea.

Note. While the above type of convergence is not explicitly defined using a metric, it is in fact equivalent to convergence in a metric on $\mathbb{M}(\mathcal{B})$, the Lévy-Prokhorov metric.

Weak convergence and convergence of measures

Definition 3.1.1. Let (μ_n) be a sequence of probability measures on $(\mathbb{R}, \mathcal{B})$, and let μ be another probability measure. We say that μ_n **converges weakly** to μ and write

$$\mu_n \xrightarrow{wk} \mu$$

if it holds for all bounded, continuous mappings $f : \mathbb{R} \to \mathbb{R}$ that $\lim_{n\to\infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu$.

Lemma 3.1.2. Let (X_n) be a sequence of random variables and let X be some other variable. Let μ be the distribution of X and let μ_n be the distribution of X_n. Then $X_n \xrightarrow{wk} X$ if and only if $\mu_n \xrightarrow{wk} \mu$.

Proof. Use that $Ef(X_n) = \int f \, \mathrm{d}\mu_n$ and $Ef(X) = \int f \, \mathrm{d}\mu$. □

Weak convergence and convergence of measures

Lemma 3.1.2 shows that in order to understand weak convergence of random variables, it suffices to consider weak convergence of probability measures.

As for the modes of convergence in Chapter 1, we may pose a series of questions about weak convergence of probability measures, such as:

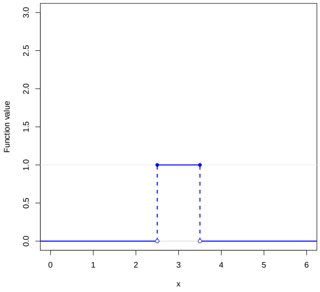
- Are weak limits unique?
- Can we identify equivalent criteria for weak convergence?
- Can we identify sufficient or necessary criteria for weak convergence?
- What kinds of stability properties does weak convergence satisfy?

Weak convergence and convergence of measures

Let $C_b^u(\mathbb{R})$ denote the set of bounded, uniformly continuous mappings $f : \mathbb{R} \to \mathbb{R}$. An important tool for the theory of weak convergence:

Lemma 3.1.3. Consider $[a, b] \subseteq (c, d)$. There exists $f \in C_b^u(\mathbb{R})$ such that $1_{[a,b]}(x) \leq f(x) \leq 1_{(c,d)}(x)$.

Graphical proof. The function $x \mapsto 1_{[a,b]}(x)$:

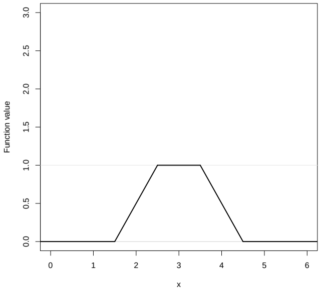


Weak convergence and convergence of measures

Let $C_b^u(\mathbb{R})$ denote the set of bounded, uniformly continuous mappings $f : \mathbb{R} \to \mathbb{R}$. An important tool for the theory of weak convergence:

Lemma 3.1.3. Consider $[a, b] \subseteq (c, d)$. There exists $f \in C_b^u(\mathbb{R})$ such that $1_{[a,b]}(x) \leq f(x) \leq 1_{(c,d)}(x)$.

Graphical proof. The function $x \mapsto f(x)$:

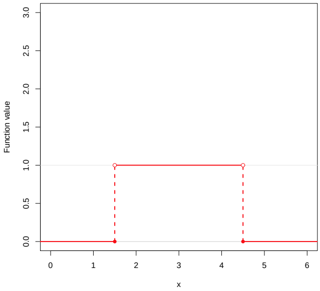


Weak convergence and convergence of measures

Let $C_b^u(\mathbb{R})$ denote the set of bounded, uniformly continuous mappings $f : \mathbb{R} \to \mathbb{R}$. An important tool for the theory of weak convergence:

Lemma 3.1.3. Consider $[a, b] \subseteq (c, d)$. There exists $f \in C_b^u(\mathbb{R})$ such that $1_{[a,b]}(x) \leq f(x) \leq 1_{(c,d)}(x)$.

Graphical proof. The function $x \mapsto 1_{(c,d)}(x)$:

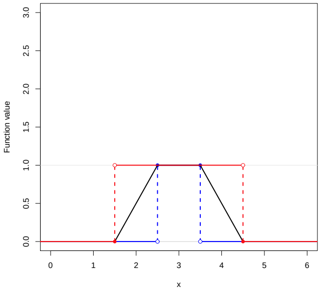


Weak convergence and convergence of measures

Let $C_b^u(\mathbb{R})$ denote the set of bounded, uniformly continuous mappings $f : \mathbb{R} \to \mathbb{R}$. An important tool for the theory of weak convergence:

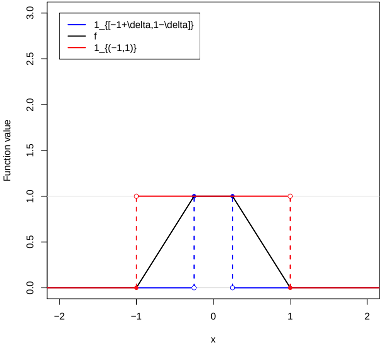
Lemma 3.1.3. Consider $[a, b] \subseteq (c, d)$. There exists $f \in C_b^u(\mathbb{R})$ such that $1_{[a,b]}(x) \leq f(x) \leq 1_{(c,d)}(x)$.

Graphical proof. All three functions:



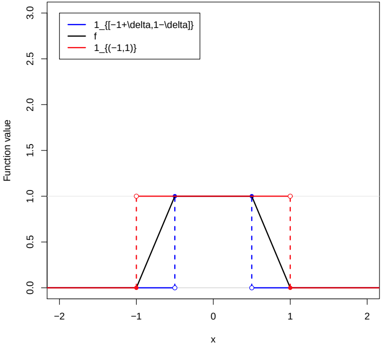
Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.750$:



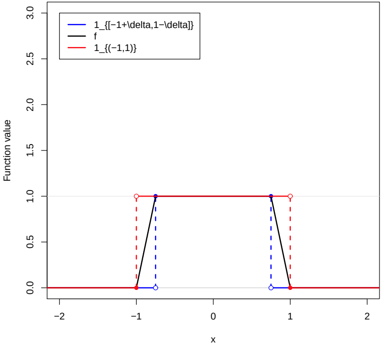
Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.500$:



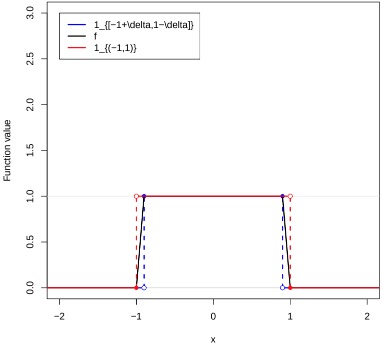
Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.250$:



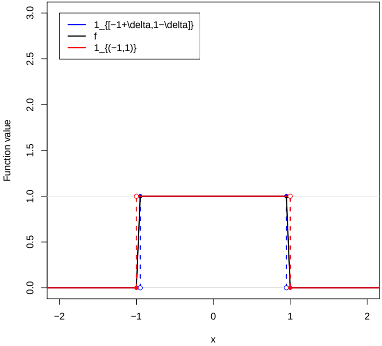
Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.010$:



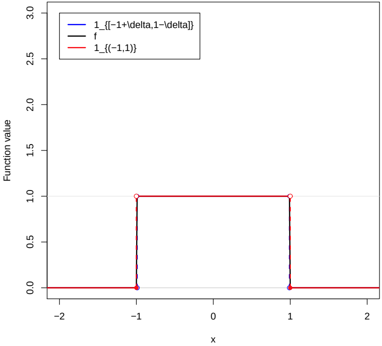
Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.050$:



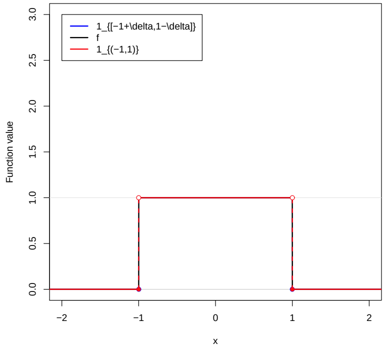
Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.010$:



Weak convergence and convergence of measures

Lemma 3.1.3 can be used to obtain uniformly continuous functions approximating discontinuous functions. Below, an illustration of how to approximate the function $x \mapsto 1_{(-1,1)}(x)$ by a function f satisfying $1_{[-1+\delta,1-\delta]}(x) \leq f(x) \leq 1_{(-1,1)}(x)$ for $\delta = 0.001$:



Weak convergence and convergence of measures

We are now ready to prove that weak limits are unique.

Lemma 3.1.4. Let μ and ν be two probability measures on $(\mathbb{R}, \mathcal{B})$. If $\int f \, \mathrm{d}\mu = \int f \, \mathrm{d}\nu$ for all $f \in C_b^u(\mathbb{R})$, then $\mu = \nu$.

Proof. Let $a < b$ and take f_n with $1_{[a+1/n, b-1/n]} \leq f \leq 1_{(a,b)}$. Note that $\lim_{n\to\infty} f_n(x) = 1_{(a,b)}$ and apply the dominated convergence theorem to obtain $\mu(a,b) = \nu(a,b)$. □

Lemma 3.1.5. Let (μ_n) be a sequence of probability measures. If $\mu_n \xrightarrow{wk} \mu$ and $\mu_n \xrightarrow{wk} \nu$, then $\mu = \nu$.

Proof. Apply Lemma 3.1.4. □

Weak convergence and convergence of measures

Next, we wish to prove that in order to show weak convergence, if suffices to prove $\lim_{n\to\infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu$ for $f \in C_b^u(\mathbb{R})$ instead of $f \in C_b(\mathbb{R})$.

Lemma 3.1.6. Let (μ_n) be a sequence of probability measures, and let μ be some other probability measure. Assume that for all $f \in C_b^u(\mathbb{R})$, $\lim_{n\to\infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu$. Then

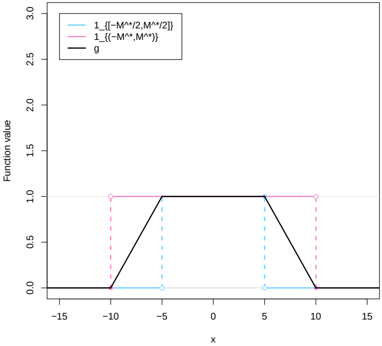
$$\lim_{M\to\infty} \sup_{n\geq 1} \mu_n([-M, M]^c) = 0.$$

Sequences satisfying this limit relationship are said to be **tight**.

Weak convergence and convergence of measures

Proof. Let $\varepsilon > 0$. Take $M^* > 0$ such that $\mu([-M^*/2, M^*/2]^c) < \varepsilon$.

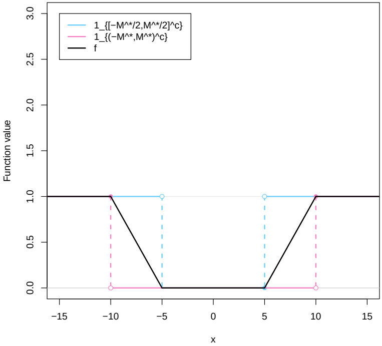
Take $g \in C_b^u(\mathbb{R})$ with $1_{[-M^*/2, M^*/2]}(x) \le g(x) \le 1_{(-M^*, M^*)}(x)$.



Weak convergence and convergence of measures

Proof. Let $\varepsilon > 0$. Take $M^* > 0$ such that $\mu([-M^*/2, M^*/2]^c) < \varepsilon$.

Define $f = 1 - g$ and obtain $1_{(-M^*,M^*)^c}(x) \leq f(x) \leq 1_{[-M^*/2,M^*/2]^c}(x)$.



Weak convergence and convergence of measures

Theorem 3.1.7. Let (μ_n) be a sequence of probability measures on $(\mathbb{R}, \mathcal{B})$, and let μ be some other probability measure. Then $\mu_n \xrightarrow{wk} \mu$ if and only if $\lim_{n\to\infty} \int f \,\mathrm{d}\mu_n = \int f \,\mathrm{d}\mu$ for $f \in C_b^u(\mathbb{R})$.

Proof. Assume that the criterion holds. We need to take $f \in C_b$ and show that $\lim_{n\to\infty} \int f \,\mathrm{d}\mu_n = \int f \,\mathrm{d}\mu$. Fix $\varepsilon > 0$. Using Lemma 3.1.6, take $M > 0$ so that $\mu_n([-M, M]^c) \leq \varepsilon$ for $n \geq 1$ and $\mu([-M, M]^c) \leq \varepsilon$.

For $h \in C_b(\mathbb{R})$ with $h(x) = f(x)$ for $x \in [-M, M]$ and $\|h\|_\infty \leq \|f\|_\infty$,

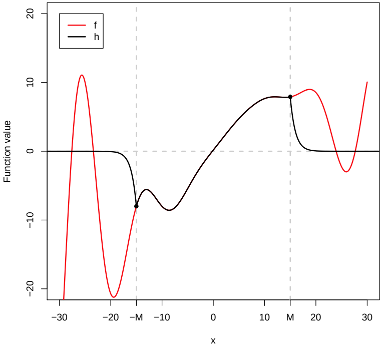
$$\left| \int f \,\mathrm{d}\mu - \int h \,\mathrm{d}\mu \right| \leq 2\varepsilon \|f\|_\infty,$$

and similarly for μ_n, $n \geq 1$.

Weak convergence and convergence of measures

Proof (continued). Define h by

$$h(x) = \begin{cases} f(-M)\exp(M+x) & \text{for } x < -M \\ f(x) & \text{for } -M \le x \le M \\ f(M)\exp(M-x) & \text{for } x > M \end{cases}$$



Weak convergence and convergence of measures

Proof (continued). We have $h(x) = f(x)$ for $x \in [-M, M]$ and $\|h\|_\infty \le \|f\|_\infty$. By the triangle inquality, we obtainy

$$\left| \int f \, \mathrm{d}\mu_n - \int f \, \mathrm{d}\mu \right| \le 4\varepsilon \|f\|_\infty + \left| \int h \, \mathrm{d}\mu_n - \int h \, \mathrm{d}\mu \right|,$$

and as $h \in C_b^u(\mathbb{R})$, the result follows. $\square$

Weak convergence and convergence of measures

Two properties of weak convergence:

Lemma 3.1.8. Let $h : \mathbb{R} \to \mathbb{R}$ be continuous. If $\mu_n \xrightarrow{wk} \mu$, then $h(\mu_n) \xrightarrow{wk} h(\mu)$.

Lemma 3.1.9 (Scheffé). Assume that there is a measure ν such that $\mu_n = g_n \cdot \nu$ for $n \geq 1$ and $\mu = g \cdot \nu$. If $\lim_{n\to\infty} g_n(x) = g(x)$ ν-almost surely, then $\mu_n \xrightarrow{wk} \mu$.

Proof. Show that $\int |g_n - g| \,\mathrm{d}\nu = 2 \int (g_n - g)^- \,\mathrm{d}\nu$ and apply the dominated convergence theorem to obtain $\lim_{n\to\infty} \int |g_n - g| \,\mathrm{d}\nu$. Use this to obtain the result. $\square$

Weak convergence and convergence of measures

Example 3.1.10. Let (x_n) be a sequence in $\mathbb{R}$ converging to x. It then holds that $\varepsilon_{x_n} \xrightarrow{wk} \varepsilon_x$. $\circ$

Example 3.1.11. Let μ_n be the uniform distribution on $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\}$. Then μ_n converges weakly to the uniform distribution on $[0, 1]$. $\circ$

Example 3.1.12. Let (ξ_n) and (σ_n) be sequences in $\mathbb{R}$ with limits ξ and σ, respectively, where $\sigma > 0$. Let μ_n be the normal distribution with mean ξ_n and variance σ_n^2. Then μ_n converges weakly to the normal distribution with mean ξ and variance σ^2. $\circ$