HOMEWORK: CONVERGENCE OF RANDOM VARIABLES AND CHARACTERISTIC FUNCTIONS

1. Lévy's criterion of convergence in law

In this part, $(X_n)_{n\in\mathbb{N}}$ is a sequence of real random variables; their laws in $\mathscr{M}^1(\mathbb{R})$ are denoted μ_n, and their characteristic functions are denoted $\phi_n(t) = \mu_n(\mathrm{e}^{\mathrm{i}tx}) = \mathbb{E}[\mathrm{e}^{\mathrm{i}tX_n}]$.

(1) Show that for any real valued random variable X, the characteristic function $\phi(t) = \mathbb{E}[\mathrm{e}^{\mathrm{i}tX}]$ is a continuous function $\mathbb{R} \to \mathbb{C}$.

In the following we suppose that $(\phi_n)_{n\in\mathbb{N}}$ converges pointwise to a function $\phi : \mathbb{R} \to \mathbb{C}$ which is continuous at $t = 0$:

$$\forall t \in \mathbb{R}, \quad \lim_{n\to\infty} \phi_n(t) = \phi(t); \tag{1}$$

$$\text{and} \quad \lim_{t\to 0} \phi(t) = \phi(0) = 1. \tag{2}$$

Questions 2. to 4. deal with the tightness of the sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ under these hypotheses; Questions 5. to 8. deal with the unicity of a limit of a subsequence.

(2) Using Fubini's theorem, show that for any law $\mu \in \mathscr{M}^1(\mathbb{R})$ of characteristic function ϕ, and any $\varepsilon > 0$,

$$I_\varepsilon = \int_{-\varepsilon}^{\varepsilon} (1 - \phi(t))\, dt = 2\varepsilon \int_{\mathbb{R}} \left(1 - \frac{\sin(\varepsilon x)}{\varepsilon x}\right) \mu(dx).$$

Show that if $|\varepsilon x| \geq 2$, then $1 - \frac{\sin(\varepsilon x)}{\varepsilon x} \geq \frac{1}{2}$. Deduce from it the inequality

$$\mu(\{x,\ |\varepsilon x| \geq 2\}) \leq \frac{I_\varepsilon}{\varepsilon}.$$

(3) Fix $\eta > 0$. Under the hypotheses (1) and (2), show that there is an $\varepsilon > 0$ and an integer N such that

$$\forall n \geq N,\ \frac{I_{\varepsilon,n}}{\varepsilon} \leq \eta.$$

Conclude that (1) and (2) imply the tightness of the sequence of laws $(\mu_n)_{n\in\mathbb{N}}$.

(4) Show that ϕ is the characteristic function of a law $\mu \in \mathscr{M}^1(\mathbb{R})$ (hint: use a convergent subsequence of $(\mu_n)_{n\in\mathbb{N}}$).

(5) Let μ and ν be two probability measures with characteristic functions ϕ_μ and ϕ_ν. Prove the Parseval identity

$$\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx}\, \phi_\mu(x)\, \nu(dx) = \int_{\mathbb{R}} \phi_\nu(y - t)\, \mu(dy). \tag{3}$$

Date: February 27, 2013.

(6) Take $\nu = \mathcal{N}(0,\varepsilon)$, a Gaussian law of variance ε with density and characteristic function

$$\nu(dx) = \frac{1}{\sqrt{2\pi\varepsilon}}\,\mathrm{e}^{-\frac{x^2}{2\varepsilon}}\,dx \qquad ; \qquad \phi_\nu(t) = \mathrm{e}^{-\frac{\varepsilon t^2}{2}}.$$

Denote X a random variable under the law μ, and $Y_{\varepsilon^{-1}}$ an independent random variable under the law $\mathcal{N}(0,\varepsilon^{-1})$. Prove that the quantity of Equation (3), viewed as a function $D(\varepsilon,t)$ of ε and t, is proportional to the density (in t) of the law of $X+Y_{\varepsilon^{-1}}$.

(7) Verify that $Y_{\varepsilon^{-1}}$ converges in probability to the constant 0 as ε goes to infinity, and that $X+Y_{\varepsilon^{-1}} \rightharpoonup X$.

(8) Show that if ϕ_μ is known, then so is the law μ, so that $\mu \mapsto \phi_\mu$ is injective.

(9) Conclude that under the hypotheses (1) and (2), the laws μ_n converge to a law μ of a random variable (for the topology of convergence in law in $\mathscr{M}^1(\mathbb{R})$). This is Lévy's continuity theorem.

(10) Show the converse implication: if $\mu_n \rightharpoonup \mu$, then (1) and (2) hold with $\phi_\mu = \phi$.

(11) Application: let B_n be a sequence of independent random Bernoulli variables with $\mathbb{P}[B_n = 1] = \frac{1}{n}$ and $\mathbb{P}[B_n = 0] = 1 - \frac{1}{n}$. Recall that $\sum_{k=1}^n \frac{1}{k} = \log n + O(1)$. We set

$$\widetilde{X}_n = \frac{\sum_{k=1}^n \left(B_k - \frac{1}{k}\right)}{\sqrt{\log n}} \qquad ; \qquad X_n = \frac{\left(\sum_{k=1}^n B_k\right) - \log n}{\sqrt{\log n}}.$$

Show that

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}t\widetilde{X}_n}] = \prod_{k=1}^n \left(1 - \frac{t^2}{2k\log n} + O\left(\frac{t^2}{k^2\,\log n} + \frac{t^3}{k\,(\log n)^{3/2}}\right)\right),$$

with a $O(\cdot)$ uniform in k. Conclude that X_n converges in law to a standard Gaussian variable.

Notice that X_n is a model for the number of disjoint cycles of a random permutation of n elements.

2. Mod-convergence and Berry-Esseen estimates

Lévy's criterion is mostly used in order to prove convergence towards a Gaussian random variable. In this second part we measure the difference between the distribution of the X_n's and the Gaussian distribution under slightly stronger hypotheses than before. Hence, we consider a sequence of real-valued random variables $(Y_n)_{n\in\mathbb{N}}$ such that

$$\forall n\in\mathbb{N},\ \ \forall t\in\mathbb{R},\ \ \mathbb{E}[\mathrm{e}^{\mathrm{i}tY_n}] = \phi_n(t) = \mathrm{e}^{-\frac{\lambda_n t^2}{2}}\,\psi_n(t); \tag{4}$$

$$\lambda_n \geq 0,\ \ \lim_{n\to\infty} \lambda_n = +\infty; \tag{5}$$

$$\lim_{n\to\infty} \psi_n = \psi \quad \text{at speed } o\left(\frac{1}{\sqrt{\lambda_n}}\right). \tag{6}$$

In Equation (4), we call the functions $\psi_n(t)$ the *residues* of the characteristic functions $\phi_n(t)$. We assume that they are continuously differentiable functions on the real line, and that ψ_n converge uniformly and sufficiently fast on every compact to the function ψ, which is itself continuously differentiable (and with $\psi_n(0) = \psi(0) = 1$); this is the meaning of Equation (6). So,

$$\forall \varepsilon > 0,\ \ \forall T > 0,\ \ \exists N,\ \ \forall n \geq N,\ \sup_{t \in [-T,T]} |\psi_n(t) - \psi(t)| \leq \frac{\varepsilon}{\sqrt{\lambda_n}}.$$

(1) Set $X_n = Y_n/\sqrt{\lambda_n}$. Show that X_n converges in law to a standard Gaussian variable of mean 0 and variance 1. For this reason, Hypotheses (4)-(6) are called hypotheses of *mod-Gaussian convergence.*

(2) For two probability measures μ and ν on $\mathbb{R}$, one defines their Kolmogorov distance as

$$d(\mu, \nu) = \sup_{x \in \mathbb{R}} |F_\mu(x) - F_\nu(x)|,$$

where $F_\mu(x)$ is the cumulative distribution function of μ, that is $\mu(-\infty, x)$. Prove that if μ is absolutely continuous with respect to the Lebesgue measure, then

$$\mu_n \rightharpoonup \mu \iff d(\mu_n, \mu) \to 0.$$

One can use freely Dini's theorem, which says that bounded increasing functions that converge pointwise converge in fact uniformly (optional: prove Dini's theorem).

Hence, in the following, we shall measure the convergence $X_n \rightharpoonup \mathcal{N}(0,1))$ by computing $d(\mu_{Y_n}, \mathcal{N}_{(0,\lambda_n)}) = d(\mu_{X_n}, \mathcal{N}_{(0,1)})$. For $T > 0$, we set

$$\Delta_T(x) = \frac{1 - \cos(Tx)}{\pi\, T x^2}.$$

Questions 3. to 5. are devoted to a proof of Berry's lemma, which relates the Kolmogorov distance to the behavior of characteristic functions. Then, in Questions 6. to 8., we apply this lemma to the situation of mod-convergence.

(3) Fix $T > 0$. Show that $\Delta_T(x) \geq 0$ for all $x \in \mathbb{R}$; that $\int_{\mathbb{R}} \Delta_T(x)\, dx = 1$; and that

$$\widehat{\Delta_T}(t) = \int_{\mathbb{R}} \Delta_T(x)\, \mathrm{e}^{\mathrm{i}tx}\, dx = \begin{cases} 1 - \frac{|t|}{T} & \text{if } |t| \leq T, \\ 0 & \text{otherwise.} \end{cases}$$

Show also that the probability measure $\Delta_T(x)\, dx$ gives to the set $\{x,\ \ |x| \geq h\}$ a mass smaller than $\frac{4}{\pi T h}$.

(4) Let F be the cumulative distribution function of a probability measure, and G be a bounded function with

$$\lim_{x \to -\infty} G(x) = 0 \quad ; \quad \lim_{x \to \infty} G(x) = 1 \quad ; \quad |G'(x)| \leq m$$

for a certain constant m. We denote

$$D(x) = F(x) - G(x) \quad ; \quad D_T(x) = \int_{\mathbb{R}} D(x-y)\, \Delta_T(y)\, dy;$$

$$\eta = \sup_{x \in \mathbb{R}} |D(x)| \quad ; \quad \eta_T = \sup_{x \in \mathbb{R}} |D_T(x)|.$$

If $\eta = 0$, show that $\eta_T = 0$. Otherwise, we fix an element x_0 such that $|D(x_0)| = \eta$; for instance we assume $D(x_0) = \eta$. Show that if $h = \frac{\eta}{2m}$ and $x = x_0 + h$, then

$$D(x - y) \geq \frac{\eta}{2} + my \quad \text{for all } |y| \leq h.$$

Prove then that

$$\eta_T \geq \frac{\eta}{2} - \frac{12m}{\pi T}$$

(distinguish the cases $\eta = 0$ and $\eta > 0$, and in this case split the integral $D_T(x)$ in two parts).

(5) Suppose that $F = F_\mu$ and G have characteristic functions

$$\phi_\mu(t) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\,\mu(dx) \quad ; \quad \phi_G(t) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\,g(x)\,dx = \widehat{g}(t) \quad \text{with } g = G'.$$

Using Fourier inversion formula $f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(t)\,\mathrm{e}^{-\mathrm{i}tx}\,dt$, prove that

$$D_T(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\Delta_T}(t)\,\left(\frac{\phi_\mu(t)-\phi_G(t)}{-\mathrm{i}t}\right)\mathrm{e}^{-\mathrm{i}tx}\,dt;$$

$$\eta \leq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\phi_\mu(t)-\phi_G(t)}{t}\right|\,dt \;+\frac{24m}{\pi T}. \tag{7}$$

This last inequality is *Berry's lemma.*

(6) Under the hypotheses (4)-(6), show that

$$\phi_{\mu_n}(t) = \mathbb{E}[\mathrm{e}^{\mathrm{i}tX_n}] = \mathrm{e}^{-\frac{t^2}{2}}\left(1+\frac{\psi'(0)\,t}{\sqrt{\lambda_n}} + o\!\left(\frac{t}{\sqrt{\lambda_n}}\right)\right).$$

Prove also that $\mathrm{e}^{-\frac{t^2}{2}}\left(1+\frac{\psi'(0)\,t}{\sqrt{\lambda_n}}\right)$ is the Fourier transform of

$$g_n(x) = \frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\left(1-\frac{\psi'(0)\,\mathrm{i}x}{\sqrt{\lambda_n}}\right).$$

We set $G_n(x) = \int_{-\infty}^{x} g_n(y)\,dy$; check that this function satisfies the previous assumptions.

(7) In the previous setting, prove that

$$\eta_n = \sup_{x\in\mathbb{R}} |F_{\mu_n}(x) - G_n(x)| = o\!\left(\frac{1}{\sqrt{\lambda_n}}\right)$$

(hint: take $T = K\sqrt{\lambda_n}$ with K big, and split the integral in (7) in two parts

$$t \in [-\varepsilon\sqrt{\lambda_n}, \varepsilon\sqrt{\lambda_n}];$$
$$t \in [-K\sqrt{\lambda_n}, K\sqrt{\lambda_n}] \setminus [-\varepsilon\sqrt{\lambda_n}, \varepsilon\sqrt{\lambda_n}].$$

with ε sufficiently small).

(8) Assuming $\psi'(0) \neq 0$, show that

$$d(\mu_{X_n}, \mathcal{N}_{(0,1)}) = \frac{|\mathrm{Im}(\psi'(0))|\,(1+o(1))}{\sqrt{2\pi\lambda_n}}.$$

(9) Optional: apply this result to the example of 1.(11).