

3. INFINITELY DIVISIBLE DISTRIBUTIONS

GIOVANNI PISTONE

The following lecture notes are based on Sasvári [2] and Sato [3].

1. **Definition** (Convolution). *Let X, Y be independent random variables with values in \mathbb{R}^n and distributions μ_X, μ_Y , respectively. The convolution $\mu_X * \mu_Y$ is the distribution of $X + Y$, that is for all bounded $f: \mathbb{R}^n \rightarrow \mathbb{R}$*

$$\int f(z) (\mu_X * \mu_Y)(dz) = \mathbb{E}(f(X + Y)) = \iint f(x + y) \mu_X(dx) \mu_Y(dy).$$

2. **Proposition** (Convolution).

(1) *If μ_1 has density p_1 with respect to the Lebesgue measure, then $\mu_1 * \mu_2$ has density $p_1 * \mu_2$ given by*

$$p_1 * \mu_2(z) = \int p_1(z - y) \mu_2(dy).$$

(2) *If both have densities p_1, p_2 , respectively, with respect to the Lebesgue measure, then $\mu_1 * \mu_2$ has density $p_1 * p_2$ given by*

$$(p_1 * p_2)(z) = \int p_1(z - y)p_2(y) dy = \int p_1(x)p_2(z - x) dx.$$

(3) *If the measurable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, and μ is a measure, define $f * \mu(z) = \int f(z - x) \mu(dx)$ if the integral exists a.e, i.e. $\int |f(z - x)| \mu(dx) < \infty$ a.s in z . If the measure μ is finite and $f \in L^a(dx)$, $1 \leq a \leq \infty$, then $f * \mu$ exists and $\|f * \mu\|_a \leq \|f\|_a$.*

Exercise.

(1) For each bounded $f: \mathbb{R}^n \rightarrow \mathbb{R}$

$$\begin{aligned} \int f(z) (\mu_1 * \mu_2)(dz) &= \iint f(x + y) p_1(x) dx \mu_2(dy) = \\ &= \int \left(\int f(x + y) p_1(x) dy \right) \mu_2(dy) = \int \left(\int f(z) p_1(z - x) dz \right) \mu_2(dy) = \\ &= \int f(z) \left(\int p_1(z - y) \mu_2(dy) \right) dz. \end{aligned}$$

(2) For each bounded $f: \mathbb{R}^n \rightarrow \mathbb{R}$

$$\begin{aligned} \int f(z) (\mu_1 * \mu_2)(dz) &= \iint f(x+y) p_1(x) dx p_2(y) dy = \\ &= \int \left(\int f(x+y) p_2(y) dy \right) p_1(x) dx = \int \left(\int f(z) p_2(z-x) dz \right) p_1(x) dx = \\ &= \int f(z) \left(\int p_1(x) p_2(z-x) dx \right) dz. \end{aligned}$$

(3) From $\iint |f(z-x)|^a dx \mu(dz) = \int |f(y)|^a dy$.

□

3. Exercise.

- (1) Compute $\chi * f$ for $\chi = \delta_a$ and $f \in C_b$.
- (2) Compute $\chi * f$ for $\chi = (\delta_a - \delta_b)^{*n}$, $n = 1, 2, 3$ and $f \in C_b$. If f is a polynomial of degree $n-1$, then $(\delta_a - \delta_b)^{*n} * f = 0$.
- (3) Compute $\chi * f$ for $\chi(dx) = A^{-1}(0 \leq x \leq A)dx$. If χ is a complex measure, define $\tilde{\chi}(B) = \overline{\chi(-B)}$. Compute $\chi * \tilde{\chi}$ and $\chi * \tilde{\chi} * f$, $f \in C_b$.
- (4) Let $\mu_\sigma = N(0, \sigma^2 I)$, and $f \in C_b$. Then $\lim_{\sigma \rightarrow 0} f * \mu_\sigma(z) = f(z)$.

Solution.

- (1) $\delta_a * f(y) = \int f(y-x) \delta_a(dx) = f(y-a)$ is the translation of f .
- (2) $(\delta_a - \delta_b) * f(y) = f(y-a) - f(y-b)$. If f is constant, then $f(y-a) - f(y-b) = 0$. $(\delta_a - \delta_b)^{*2} * f(y) = f(y-2a) - 2f(y-a-b) + f(y-2b)$. If $f = u^t y + v$ is a polynomial of degree 1, then $f(y-a) - f(y-b) = u^t(y-a) - u^t(y-b) = u^t(b-a)$ is constant. In general, if f is a polynomial of degree n , then $(\delta_a - \delta_b) * f(y) = f(y-a) - f(y-b) = \sum_{k=1}^n \frac{1}{k!} f^{(k)}(y-b)(b-a)^k$ is a polynomial of degree at most $n-1$, e.g. $(y-a)^2 - (y-b)^2 = a(y-b)(b-a) + (a-b)^2$.
- (3) Continuity and dominated convergence:

$$\begin{aligned} f * \mu_\sigma(y) &= \int f(y-x) (2\pi)^{-n/2} \sigma^{-n} e^{-\|x\|^2/2\sigma^2} dy = \\ &= \int f(y-\sigma z) (2\pi)^{-n/2} e^{-\|x\|^2/2} dz \rightarrow \int f(y) (2\pi)^{-n/2} e^{-\|x\|^2/2} dz = f(y). \end{aligned}$$

□

4. Definition (Characteristic function, inverse Fourier transform).

- (1) Let X be a random variable in \mathbb{R}^n , with distribution μ_X . The characteristic function of X is the function $\phi_X: \mathbb{R}^n \rightarrow \mathbb{C}$ defined by

$$\phi_X(t) = E(e^{i\langle t, X \rangle}) = \int e^{i\langle t, x \rangle} \mu(dx), \quad i = \sqrt{-1}.$$

- (2) The function $\check{\mu}_X(t) = \int e^{i\langle t, x \rangle} \mu(dx)$ is also called inverse Fourier transform of μ . If μ has a density f with respect to the Lebesgue measure, then the inverse Fourier transform of f is $\check{f}(t) = \int e^{i\langle t, x \rangle} f(x) dx = \check{\mu}(t)$. If $f \in L^1(dx)$ then the Fourier transform $\hat{f}(t) = \int e^{-i\langle t, x \rangle} f(x) dx$ is defined for all t and $\|\hat{f}\|_1 \leq \|f\|_1$.

5. Definition (Positive definite). A function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ is positive definite if for all $m \in \mathbb{N}$, $x_1, \dots, x_m \in \mathbb{R}^n$, $c_1, \dots, c_m \in \mathbb{C}$ it holds $\sum_{i,j=1}^m c_i \bar{c}_j f(x_i - x_j) \geq 0$. In other words, for all $m \in \mathbb{N}$, $x_1, \dots, x_m \in \mathbb{R}^n$ the matrix $A = [f(x_i - x_j)]_{i,j=1}^m$ is positive definite, that is for all $u \in \mathbb{C}^m$, it holds $u^* A u \geq 0$, $u^* = \overline{u^t}$.

6. **Definition** (Hermitian). A function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ is Hermitian if $f(-t) = \overline{f(t)}$, $t \in \mathbb{R}^n$.

7. **Proposition** (Properties of positive definite functions).

- (1) If f is positive definite, then f is Hermitian and $f(0) \geq 0$.
- (2) For a complex measure μ , define $\tilde{\mu}$ by $\int f(x) \tilde{\mu}(dx) = \int \overline{f(-x)} \mu(dx)$. Analogously, for each $g: \mathbb{R}^n \rightarrow \mathbb{C}$, define $\tilde{g}(t) = \overline{g(-t)}$. If f is continuous and positive definite, then $\mu * \tilde{\mu} * f$ is positive definite. Similarly for $g * \tilde{g} * f(t)$.
- (3) If $g \in L^2(\mathbb{R}^n, \lambda; \mathbb{C})$, that is $\int |g(x)|^2 dx < \infty$, then $g * \tilde{g}$ exists and is positive definite.
- (4) Let X_t , $t \in \mathbb{R}^n$ be a family of random variables such that $\text{Cov}(X_t, X_s) = \rho(t - s)$. Then ρ is positive definite.
- (5) If f is positive definite, then $x \mapsto e^{i\langle t, x \rangle} f(x)$ is positive definite.
- (6) If f is positive definite and integrable, the $\int f(x) dx \geq 0$.

Exercise.

- (1) Take $m = 1$ $c_1 = 1$, $x_1 = 0$: $c_1 \overline{c_1} f(0 - 0) = f(0) \geq 0$. Take $m = 2$, $x_1 = 0$, $x_2 = x$: $|c_1|^2 f(0) + c_1 \overline{c_2} f(-x) + c_2 \overline{c_1} f(x) + |c_2|^2 f(0) \geq 0$. In particular, with $c_1 = c_2 = 1$, we have $f(-x) + f(x) \in \mathbb{R}$, and, with $c_1 = 1$, $c_2 = i$, we have $i(f(x) - f(-x)) \in \mathbb{R}$. It follows $f(x) = [(f(-x) + f(x)) - i \cdot i(f(x) - f(-x))]/2$ and $f(-x) = [(f(-x) + f(x)) + i \cdot i(f(x) - f(-x))]/2$.
- (2) It is enough to consider $\mu = \sum_{i=1}^m c_i \delta_{x_i}$, $\tilde{\mu} = \sum_{j=1}^m \overline{c_j} \delta_{-x_j}$. In such a case

$$\mu * \tilde{\mu} * f(z) = \iint f(z - x - y) \mu(dx) \tilde{\mu}(dy) = \sum_{i,j=1}^m c_i \overline{c_j} f(z - x_i + x_j).$$

We check the positive definiteness with

$$\begin{aligned} \sum_{h,k=1}^M d_h \overline{d_k} \mu * \tilde{\mu} * f(z) &= \iint f(z - x - y) \mu(dx) \tilde{\mu}(dy) = \\ \sum_{h,k=1}^M d_h \overline{d_k} \sum_{i,j=1}^m c_i \overline{c_j} f(z_h - z_k - x_i + x_j) &= \sum_{h,k=1}^M \sum_{i,j=1}^m d_h \overline{d_k} c_i \overline{c_j} f((z_h - x_i) - (z_k - x_j)) = \\ &= \sum_{(h,i),(k,j)} (c_i d_h) \overline{(c_j d_k)} (f(y_{h,i} - y_{k,j})) \geq 0. \end{aligned}$$

- (3) The existence of $g * \tilde{g}(y) = \int g(y - x) \overline{g(-x)} dx$ follows from

$$|g * \tilde{g}(y)| \leq \sqrt{\int |g(y - x)|^2 dx} \sqrt{\int |\overline{g(-x)}|^2 dx} = \|g\|_2^2.$$

The positive definiteness is

$$\begin{aligned} \sum_{i,j=1}^m c_i \overline{c_j} \int g(y_i - y_j - x) \overline{g(-x)} dx &= \int \sum_{i,j=1}^m c_i \overline{c_j} g(y_i - y_j - x) \overline{g(-x)} dx = \\ &= \int \left| \sum_{i=1}^m c_i g(y_i - x) \right|^2 dx \geq 0. \end{aligned}$$

- (4) From the definition.
- (5) From the definition.

- (6) Let g_n be a sequence of triangular functions such that $f(z)g_n(z) \rightarrow f(z)$. Write $g_n = h_n * \tilde{h}_n$, h_n being uniform, and compute $\int f(x)\tau_n(x) dx$ as the value at 0 of a positive convolution. I.e. ($n=1$) let $h_n(x) = n^{-1/2}(0 \leq x \leq n)$ and define $g_n = h_n * \tilde{h}_n$. Then $g_n(0) = 1$, $0 \leq g_n(x) \leq 1$ and $g_n(x) \rightarrow 1$, $n \rightarrow \infty$. From $\int f(x)g_n(x) dx = f * g_n * \tilde{g}_n(0) \geq 0$ we obtain the result. \square

8. Proposition.

- (1) The characteristic function is uniformly continuous and $\check{\mu}(0) = 1$.
- (2) If X is a random variable in \mathbb{R}^n with characteristic function ϕ_X , for each $A \in \mathbb{R}^{m \times n}$ and $a \in \mathbb{R}^m$, the random variable $Y = a + AX$ has characteristic function $\phi_Y(s) = e^{i\langle a, s \rangle} \phi_X(A^T s)$.
- (3) If X_1 and X_2 are independent random variables with values in \mathbb{R}^{n_1} and \mathbb{R}^{n_2} , and characteristic functions ϕ_1, ϕ_2 , respectively, then $X = (X_1, X_2)$ has characteristic function $\phi_X(t_1, t_2) = \phi_1(t_1)\phi_2(t_2)$.
- (4) The characteristic function is Hermitian.
- (5) The characteristic function is positive definite.

Exercise.

- (1) We have $|e^{i\langle t+h, x \rangle} - e^{i\langle t, x \rangle}| = |e^{i\langle t, x \rangle}(e^{i\langle h, x \rangle} - 1)| = |e^{i\langle h, x \rangle} - 1| \leq 2$, and $\lim_{h \rightarrow 0} O(h) = \int |e^{i\langle h, x \rangle} - 1| \mu(dx) = 0$ by dominated convergence.
- (2) From $\langle s, a + Ax \rangle = \langle s, a \rangle + \langle A^T s, x \rangle$.
- (3) From independence and $\langle (t_1, t_2), (x_1, x_2) \rangle_{n_1+n_2} = \langle t_1, x_1 \rangle_{n_1} + \langle t_2, x_2 \rangle_{n_2}$.
- (4) $\check{\mu}(-t) = \int e^{i\langle -t, x \rangle} \mu(dx) = \int \overline{e^{i\langle t, x \rangle}} \mu(dx) = \overline{\check{\mu}(t)}$.
- (5)

$$\begin{aligned} \sum_{i,j=1}^n c_i \overline{c_j} \check{\mu}(t_1 - t_j) &= \int \sum_{i,j=1}^n c_i \overline{c_j} e^{i\langle t_i - t_j, x \rangle} \mu(dx) = \\ &= \int \sum_{i,j=1}^n c_i \overline{c_j} e^{i\langle t_i, x \rangle} \overline{e^{i\langle t_j, x \rangle}} \mu(dx) = \int \left| \sum_{i=1}^n c_i e^{i\langle t_i, x \rangle} \right|^2 \mu(dx) > 0 \end{aligned}$$

\square

The following proposition requires the use of complex logarithms, which are not easily defined because the complex exponential function $e^z = e^{\Re z} e^{i\Im z} = e^{\Im z} (\cos(\Im z) + i \sin(\Im z))$ is not invertible as $e^z = e^{z+i2\pi k}$, $k \in \mathbb{Z}$.

9. Proposition (Cumulant function). *Let $\phi: \mathbb{R}^n \rightarrow \mathbb{C}$ be the characteristic function of the probability measure μ and assume that $\phi(t) \neq 0$, $t \in \mathbb{R}^n$. There exists a unique continuous function $\psi: \mathbb{R}^n \rightarrow \mathbb{C}$ such that $\phi(t) = e^{\psi(t)}$ and $\kappa(0) = 0$. Such a function is Hermitian. It is called the cumulant function of μ .*

Proof. In steps.

- (1) The cumulant function is unique. In fact, if $\phi(t) = e^{\psi_1(t)} = e^{\psi_2(t)}$, then $\Re \psi_i(t) = \log(|\phi(t)|)$, $i = 1, 2$, hence $e^{i\Im \psi_1(t)} = e^{i\Im \psi_2(t)}$, hence $\Im \psi_i(t) - \Im \psi_j(t) = 2\pi k(t)$. The function $t \mapsto k(t)$ is continuous and integer valued on \mathbb{R}^n , then constant and equal 0 at $t = 0$.
- (2) The cumulant function is Hermitian. From $\phi(-t) = \overline{\phi(t)}$ it follows $e^{i\Im \psi(-t)} = e^{-i\Im \psi(t)}$, hence $\Im \psi(-t) + \Im \psi(t) = 2\pi k(t)$.

- (3) Continuous argument. As ϕ is never zero, we can define $f(t) = \phi(t)/|\phi(t)|$ so that $f: \mathbb{R}^n \rightarrow \mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ is continuous and $f(0) = 1$. A *continuous argument* of f is a continuous $\theta: \mathbb{R}^n \rightarrow \mathbb{R}$ such that $f(t) = e^{i\theta(t)}$ and $\theta(0) = 0$. If a continuous argument exist, then $\phi(t) = |\phi(t)| e^{i\theta(t)} = e^{\log(|\phi(t)|) + i\theta(t)} = e^{\psi(t)}$, with $\psi = \log|\phi| + i\theta$ continuous and $\psi(0) = \log 1 + \theta(0) = 0$.
- (4) Consider a continuous $f: B \rightarrow \mathbb{T}$, $B \subset \mathbb{R}^n$, such that $f(B) \neq \mathbb{T}$. Choose $e^{i\alpha} \in \mathbb{T} \setminus f(B)$. Then the function $z \mapsto \arg(e^{-i\alpha}z) + \alpha$ is a continuous bijection of $\mathbb{T} \setminus \{e^{i\alpha}\}$ onto $] \alpha, \alpha + 2\pi[$, hence $f(t) = e^{i(\arg(e^{-i\alpha}f(t)) + \alpha)}$ so that $t \mapsto \theta(t) = \arg(e^{-i\alpha}f(t)) + \alpha$ is a continuous argument of f on B .
- (5) Let $f_1, f_2: B \rightarrow \mathbb{T}$, $B \subset \mathbb{R}^n$, be continuous. Assume f_1 has a continuous argument on B , $f_1 = e^{i\theta_1}$, and $f_1(t) + f_2(t) \neq 0$, $t \in B$. Then $f_1/f_2: \mathbb{R}^n \rightarrow \mathbb{C}$ never equals -1 , hence it has a continuous argument on B , $f_1/f_2 = e^{i\theta}$, so that $f_2 = e^{i(\theta_1 - \theta)}$.
- (6) Consider the function $f^R: B(R) = \{t \mid \|t\|^2 \leq R\} \times [0, 1] \ni (t, \alpha) \mapsto f(\alpha t)$. The function f^R is continuous on a compact set, hence uniformly continuous, so that there exists a $n \in \mathbb{N}$ such that $|f^R(\alpha_1 t) - f^R(\alpha_2 t)| \leq 1$ for all t if $|\alpha_1 - \alpha_2| \leq 1/n$. Consider the sequence $f_j(t) = f^R(\frac{j}{n}t)$, $j = 0, 1, \dots, n$. We proceed by finite induction on j . if $j = 0$ then $f_0(t) = f(0) = 1$ and the continuous argument is $\theta_0 = 0$. If there exist a continuous argument on $B(R)$ for f_j , $j < n$, as $|f_{j+1}(t) - f_j(t)| \leq 1$, the relation $f_{j+1}(t) + f_j(t) = 0$ is impossible, and the previous item shows that f_{j+1} has a continuous argument on $B(R)$. Finally, note that $f_n = f^R$.
- (7) Each f has a continuous argument θ_R on $B(R)$ and $\theta_{R_1}(t) = \theta_{R_2}(t)$ for all $t \in B(\min(R_1, R_2))$ because $B(\min(R_1, R_2))$ is connected and the uniqueness argument applies. Then the global continuous argument is defined by its restrictions. \square

10. Exercise.

- (1) Let μ, ν , be probability measures on \mathbb{R}^n with characteristic function respectively $\check{\mu}, \check{\nu}$. Then

$$\int e^{i\langle t, y \rangle} \check{\mu}(t) \nu(dt) = \int \check{\nu}(x + y) \mu(dx)$$

- (2) Let $X \sim \mu$ independent of $Y \sim N(0, 1)$ and $\nu \sim X + \sigma Y$. Let g_σ be the density of σY . Then $\nu = \mu * g_\sigma$ has density

$$p_\sigma(t) = \int g_\sigma(t - x) \mu(dx) = (2\pi)^{-n/2} \sigma^n \int \check{g}_{1/\sigma} \mu(dx) = \int e^{-i\langle x, y \rangle} \check{\mu}(x) g_{1/\sigma}(x) dx.$$

Proof.

- (1) We have $e^{i\langle t, y \rangle} \check{\mu}(t) = e^{i\langle t, y \rangle} \int e^{i\langle t, y \rangle} \mu(dy) = \int e^{i\langle t, x+y \rangle} \mu(dy)$ and we can take the integral with respect to ν to get $\int (e^{i\langle t, y \rangle} \check{\mu}(t)) \nu(dt) = \iint e^{i\langle t, x+y \rangle} \mu(dy) \nu(dt) = \int \check{\nu}(x + y) \mu(dx)$.
- (2) From the properties of the Gaussian density and the previous equality. \square

11. Proposition (Inversion theorems). Let denote by μ be a probability measure on \mathbb{R}^n with inverse Fourier transform $\check{\mu}$.

- (1) For all $f: \mathbb{R}^n \rightarrow \mathbb{R}$ continuous and with bounded support, $f \in C_{00}(\mathbb{R}^n)$,

$$\int f(x) \mu(dx) = (2\pi)^{-n} \lim_{T \rightarrow \infty} \int_{-T}^T \dots \int_{-T}^T \left(\int f(s) e^{-i\langle s, t \rangle} ds \right) \check{\mu}(t) dt.$$

- (2) The mapping $\mu \mapsto \check{\mu}$ is 1-to-1.
(3) If $\check{\mu}$ is integrable, that is $\int |\check{\mu}(t)| dt < \infty$, then μ has a bounded and uniformly continuous density p with respect to the Lebesgue measure λ , and

$$p(x) = (2\pi)^{-n} \int \check{\mu}(t) e^{-i\langle t, x \rangle} dt = (2\pi)^{-n} \hat{\check{\mu}}(x).$$

Exercise.

- (1) We have

$$\begin{aligned} \int_{-T}^T \cdots \int_{-T}^T \left(\int f(s) e^{-i\langle s, t \rangle} ds \right) \check{\mu}(t) dt &= \\ \int_{-T}^T \cdots \int_{-T}^T \left(\int f(s) e^{-i\langle s, t \rangle} ds \right) \left(\int e^{i\langle t, y \rangle} \mu(dy) \right) dt &= \\ \int ds f(s) \int \mu(dy) \int_{-T}^T \cdots \int_{-T}^T dt e^{i\langle t, y-s \rangle} &= \\ 2^n \int ds f(s) \int \mu(dy) \prod_{i=1}^n \frac{\sin(T(y_i - s_i))}{y_i - s_i} &= \\ 2^n \int \left(\int f(s) \prod_{i=1}^n \frac{\sin(T(y_i - s_i))}{y_i - s_i} ds \right) \mu(dy) &= \quad T(y - s) = u \\ &= 2^n \int f(y - T^{-1}u) \prod_{i=1}^n \frac{\sin(u_i)}{u_i} du \rightarrow \pi^n \quad T \rightarrow \infty \end{aligned}$$

See another version in [4, §16.6].

- (2) Follows from the previous inversion formula and the monotone class theorem. See also a direct proof based on Ex. . in [1, 14.1]
(3) Use the approximation with the Gaussian kernel of Ex. ., see [2, Th 1.3.6]

□

12. Definition (Weak convergence, convergence in distribution).

- (1) A sequence $(\mu_n)_{n \in \mathbb{N}}$ of probability measure on \mathbb{R}^n converges weakly to a probability measure μ if for all bounded and continuous $f: \mathbb{R}^n \rightarrow \mathbb{R}$ we have

$$\lim_{n \rightarrow \infty} \int f(x) \mu_n(dx) = \int f(x) \mu(dx), \quad f \in C_b(\mathbb{R}^n).$$

- (2) If $(X_n)_{n \in \mathbb{N}}$ and X are random variables in \mathbb{R}^n we say that $\lim_{n \rightarrow \infty} X_n = X$ in distribution if $\lim_{n \rightarrow \infty} \mathbb{E}(f(X_n)) = \mathbb{E}(f(X))$, $f \in C_b(\mathbb{R}^n)$.

13. Proposition. Let $\mu, \nu, \mu_n, n \in \mathbb{N}$, be probability measures on \mathbb{R}^n . If $\lim_{n \rightarrow \infty} \mu_n = \mu$ weakly, then $\lim_{n \rightarrow \infty} \nu * \mu_n = \nu * \mu$ weakly.

Proof. If $f \in C_b(\mathbb{R}^n)$, then for all x we have $(y \mapsto f(x+y)) \in C_b(\mathbb{R}^n)$, and dominated convergence implies that

$$\begin{aligned} \lim_{n \rightarrow \infty} \int f(z) (\nu * \mu_n)(dz) &= \lim_{n \rightarrow \infty} \iint f(x+y) \nu(dx) \mu_n(dy) = \\ \lim_{n \rightarrow \infty} \int \left(\int f(x+y) \mu_n(dy) \right) \nu(dx) &= \int \left(\lim_{n \rightarrow \infty} \int f(x+y) \mu_n(dy) \right) \nu(dx) = \\ &= \int \left(\int f(x+y) \mu(dy) \right) \nu(dx) = \int f(z) (\nu * \mu)(dz). \end{aligned}$$

□

14. Proposition (Lévy continuity theorem). *Let $(\phi_n)_{n \in \mathbb{N}}$ the the sequence of characteristic functions of the sequence of probability measures $(\mu_n)_{n \in \mathbb{Z}}$. If there exist the pointwise limit $\phi(t) = \lim_{n \rightarrow \infty} \phi_n(t)$, $t \in \mathbb{R}^n$, and the limit function ϕ is continuous at 0, then ϕ is a characteristic function of a probability measure μ . In such a case, for all bounded continuous function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, we have $\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu$, that is the sequence $(\mu_n)_{n \in \mathbb{N}}$ weakly converges to μ .*

Proof. See [4, 18.1] or [1, Th. 19.1].

□

15. Proposition. *The mapping $\mu \mapsto \check{\mu}$ is 1-to-1 from probability measure to positive definite functions whose value is 1 at 0.*

Proof.

□

16. Proposition (Bochner theorem). *If the function $\phi: \mathbb{R}^n \rightarrow \mathbb{C}$ is continuous, positive definite, and such that $\phi(0) = 1$, then there exists a probability measure μ such that $\phi = \check{\mu}$.*

Exercise. Assume first that ϕ is integrable. Let g_σ be the density of the Gaussian $N(0, \sigma^2 I)$ with characteristic function \check{g}_σ ,

$$g_\sigma(x) = (2\pi\sigma^2)^{-n/2} e^{-\|x\|^2/2\sigma^2}, \quad \check{g}_\sigma(t) = e^{-\sigma^2\|t\|^2/2}.$$

Note that $\check{g}_\sigma(t) = (2\pi)^{-n/2} g_{1/\sigma}(t)$. Let us compute

$$\begin{aligned} \phi * \check{g}_\sigma(t) &= \int \phi(s) \check{g}_\sigma(t-s) ds \\ &= \int \phi(s) \left(\int e^{i\langle t-s, x \rangle} g_\sigma(x) dx \right) ds \\ &= \int e^{i\langle t, x \rangle} g_\sigma(x) \left(\int e^{-i\langle s, x \rangle} \phi(s) ds \right) dx \\ &= \int e^{i\langle t, x \rangle} g_\sigma(x) \hat{\phi}(x) dx \end{aligned}$$

As $\hat{\phi}$ is nonnegative being the Fourier transform of a positive definite function, and $\hat{\phi} \leq \int |\phi|$, we can renormalize $g_\sigma \cdot \hat{\phi}$ to get a probability density, so that

$$\frac{\phi * \check{g}_\sigma(t)}{\phi * \check{g}_\sigma(0)} = \frac{\phi * g_{1/\sigma}(t)}{\phi * g_{1/\sigma}(0)}$$

is a characteristic function. As $\sigma \rightarrow 0$, letting $\sigma s = u$,

$$\phi * \check{g}_\sigma(t) = \int \phi(t-s) e^{-\sigma^2\|s\|^2/2} ds = \sigma^{-1} \int \phi(t - \sigma^{-1}u) e^{-\|u\|^2/2} du,$$

$$\frac{\phi * \check{g}_\sigma(t)}{\phi * \check{g}_\sigma(0)} = \frac{\int \phi(t - \sigma^{-1}u) e^{-\|u\|^2/2} du}{\int \phi(-\sigma^{-1}u) e^{-\|u\|^2/2} du} \rightarrow \frac{(2\pi)^{-n/2} f(t)}{(2\pi)^{-n/2} f(0)} = f(t).$$

If ϕ is not integrable, for each n the function $\phi_n: t \mapsto \phi(t)\check{g}_{1/n}(t)$ are positive definite, integrable, 1 at 0, and $\lim_{n \rightarrow \infty} \phi_n(t) = \phi(t)$. \square

17. Definition (Infinite divisibility).

- A random variable X is infinitely divisible if for all $n \in \mathbb{N}$ there exist IID random variables X_1, \dots, X_n such that $X \sim X_1 + \dots + X_n$.
- Equivalently, a probability measure μ is infinitely divisible if for all $n \in \mathbb{N}$ there exists a probability measure μ_n such that $\mu = (\mu_n)^{*n}$.
- Equivalently, a characteristic function ϕ is infinitely divisible if for all $n \in \mathbb{N}$ there exists a characteristic function ϕ_n such that $\phi = (\phi_n)^n$.

18. Proposition.

- (1) If the characteristic functions ϕ, ϕ_1 are infinitely divisible, then $\bar{\phi}, |\phi|^2, \phi\phi_1$ are infinitely divisible.
- (2) Each infinite divisible characteristic function ϕ has a cumulant function, $\phi = e^\psi$.

Exercise.

- (1) If ϕ and ϕ_1 are the characteristic functions of the random variables X and X_1 , respectively, then $\bar{\phi}$ is the characteristic function of $-X$, $|\phi|^2$ of $X - X'$, X' being an independent copy of X , $\phi\phi_1$ of $X + X_1$, X and X_1 independent.
- (2) Because of Proposition 9 we want $\phi(t) \neq 0, t \in \mathbb{R}^n$. For all $n \in \mathbb{N}$, let $\phi = (\phi_n)^n$. Then $|\phi|^2 = |\phi_n|^{2n}$, hence $|\phi|^{2/n} = |\phi_n|^2$ is a characteristic function for all n . The limit $\phi_*(t) = \lim_{n \rightarrow \infty} |\phi|^{2/n}$ is $\phi_*(t) = 1$ if $\phi(t) \neq 0$ and $\phi_*(t) = 0$ if $\phi(t) = 0$. As $\phi(t) \neq 0$ in a neighborhood of 0, then ϕ_* is a characteristic function equal to 1 in a neighborhood of 0, hence it is a characteristic function, hence continuous, so that the case $\phi(t) = 0$ is impossible. \square

19. Exercise (Table of infinitely divisible characteristic functions).

Sampling. If $\phi_i, i = 1, \dots, n$, are infinitely divisible, then $\phi_1 \otimes \dots \otimes \phi_n$ is infinitely divisible.

Affine transformation. If ϕ is infinitely divisible, then $s \mapsto e^{i\langle s, \mu \rangle} \phi(A^T s)$ is infinitely divisible. If ψ is the cumulant function, the transformed cumulant is $s \mapsto i\langle s, \mu \rangle + \psi(A^T s)$.

Dirac. The Dirac distribution δ_μ has characteristic function $\phi(t) = e^{i\langle t, \mu \rangle}$ and cumulant function $\psi(t) = i\langle t, \mu \rangle$.

Poisson. If $X \sim \text{Poisson}(\lambda)$, then X has characteristic function $\phi(t) = e^{\lambda(e^{it} - 1)}$. It is infinitely divisible. The cumulant function is

$$\psi(t) = \lambda(e^{it} - 1) = \int (e^{ity} - 1) (\lambda\delta_1)(dy) = \int (e^{ity} - 1) \nu(dy)$$

Gaussian. If $X \sim \text{Normal}_n(0, I)$, then X has characteristic function $\phi(t) = e^{-\|t\|^2/2}$. It is infinitely divisible. The cumulant function is $\psi(t) = -\frac{1}{2} \|t\|^2$.

Gamma. If $X \sim \text{Gamma}(\gamma, \lambda)$, then X has characteristic function $\phi(t) = (1 - i\lambda^{-1}t)^{-\gamma}$. It is infinitely divisible with n -th root $\phi_n(t) = (1 - i\lambda^{-1}t)^{-\gamma/n}$ and $\psi(t) = \lim_{n \rightarrow \infty} n(\phi_n(t) - 1)$. Let us consider the measure

$$\nu(dy) = (\gamma y)^{-1} e^{-\lambda y} (y > 0) dy.$$

We have

$$\begin{aligned} \int (e^{ity} - 1) \nu(dy) &= \gamma^{-1} \int_0^\infty \frac{e^{ity} - 1}{y} e^{-\lambda y} dy \\ &= i\gamma^{-1} \int_0^\infty \frac{e^{ity} - 1}{iy} e^{-\lambda y} dy \\ &= i\gamma^{-1} \int_0^\infty \left(\int_0^t e^{isy} ds \right) e^{-\lambda y} dy \\ &= i\gamma^{-1} \int_0^t ds \int_0^\infty dy e^{-(\lambda - is)y} \\ &= i\gamma^{-1} \int_0^t (\lambda - is)^{-1} ds \\ &= -\gamma^{-1} \log(\lambda - is) \Big|_{s=0}^{s=t} \\ &= \log((1 - i\lambda^{-1}t)^{-\gamma}). \end{aligned}$$

Hence

$$\psi(t) = \int (e^{ity} - 1) \nu(dy).$$

Note that the measure μ is infinite, but $y \mapsto e^{ity} - 1$ is integrable, because

$$|e^{ity} - 1| = \left| i \int_0^{ty} e^{iu} du \right| \leq |ty|,$$

so that

$$\int |e^{ity} - 1| \nu(dy) \leq \gamma^{-1} |t| \int_0^\infty e^{-\lambda y} dy = |t|/\gamma\lambda.$$

An other integral form is of interest. Let $h: \mathbb{R} \rightarrow \mathbb{R}$ be bounded, continuous, asymmetric $h(-y) = -h(y)$, equal to the identity $h(y) = y$ if $|y| < 1$ and constant for $|y| \geq 1$. Then

$$\begin{aligned} \int ith(y) \nu(dy) &= it\gamma^{-1} \int_0^1 e^{-\lambda y} dy + it\gamma^{-1} \int_1^\infty y^{-1} e^{-\lambda y} dy = \\ &it(\lambda^{-1}(e^{-\lambda} - 1) + E_1(\lambda)) = it\mu \end{aligned}$$

so that

$$\psi(t) = it\mu + \int (e^{ity} - 1 - ith(y)) \nu(dy).$$

The integrand in this form is a function that equals the Taylor remainder of order 2. In fact, if $|y| < 1$, then

$$|e^{ity} - 1 - ith(y)| = |e^{ity} - 1 - ity| \leq \frac{t^2 y^2}{2} < \frac{t^2}{2}.$$

With the same function as above,

$$\int ith(y) \nu(dy) = it\lambda \int_{|y| < 1} y \sigma(dy) + it\lambda \int_{|y| \geq 1} \text{sign}(y) \sigma(dy) = it\mu,$$

so that

$$\psi(t) = it\mu + \int (e^{ity} - 1 - ith(y)) \nu(dy).$$

Compound Poisson. Let N be a Poisson process with intensity λ and $(X_n)_{n \in \mathbb{N}}$ be IID with distribution σ . Assume N and $(X_n)_{n \in \mathbb{N}}$ independent. The process defined by $Y_t = \sum_{k=1}^{N_t} X_k$ is called *compound Poisson*. The characteristic function of Y_1 is

$$\phi(t) = \mathbb{E} \left(e^{it(\sum_{k=1}^{N_1} X_k)} \right) = \sum_{n=0}^{\infty} \mathbb{E} \left(e^{it(\sum_{k=1}^n X_k)} (N_1 = n) \right) = \sum_{n=0}^{\infty} (\check{\sigma}(t))^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{\lambda((\check{\sigma}(t))^n - 1)}.$$

The cumulant function is

$$\psi(t) = \lambda((\check{\sigma}(t))^n - 1) = \int (e^{ity} - 1) (\lambda\sigma)(dy) = \int (e^{ity} - 1) \nu(dy).$$

Approximation by Compound Poisson. Let g_1 be the Gaussian density $N(0, 1)$. Then $\check{g}_1(x) = e^{-\|x\|^2/2}$ with cumulant function $\psi(x) = -\frac{1}{2}\|x\|^2$. Consider the infinite divisibility, $\frac{1}{m}\psi(x) = \frac{1}{2m}\|x\|^2$, with is the cumulant function of $e^{-\|x\|^2/2m} = \check{g}_{m^{-1/2}}(x)$, i.e. $N(0, m^{-1})$. Consider the distribution Compound Poisson $CP(m, g_{m^{-1/2}})$, i.e. the distribution of $\sum_{k=1}^N X_i$, with $N \sim \text{Poi}(m)$, $(X_k)_k$ IID $N(0, m^{-1})$. The characteristic function is

$$t \mapsto e^{m(\check{g}_{m^{-1/2}}(t) - 1)} = e^{m(e^{-m^{-1}\|x\|^2/2} - 1)},$$

and cumulant function

$$t \mapsto m(e^{-m^{-1}\|x\|^2/2} - 1) = \int (e^{i\langle t, x \rangle} - 1) (mg_{m^{-1/2}})(x) dx.$$

The left hand side converges to $-\|x\|^2/2$. Note the peculiar convergence of the right hand side, were $g_{m^{-1/2}}(x) dx$ converges weakly to δ_0 , while $mg_{m^{-1/2}}(x) dx$ does not converge weakly, as for all f integrable

$$\int f(x) mg_{m^{-1/2}}(x) dx = \int mf(m^{-1/2}x) g_1(x) dx.$$

20. Definition. A function $f: \mathbb{R}^n \rightarrow \mathbb{C}$ is conditionally positive definite if it is Hermitian and $\sum_{i,j=1}^m c_i \bar{c}_j f(t_i - t_j) \geq 0$ for all $t_1, \dots, t_m \in \mathbb{R}^n$, $c_1, \dots, c_m \in \mathbb{C}$, $\sum_{i=1}^m c_i = 0$.

21. Proposition (Lévy-Kinchin formula). Let the μ be a probability measure on \mathbb{R}^n with characteristic function $\check{\mu}$ and cumulant function ψ , $\check{\mu}(t) = e^{\psi(t)}$. The following conditions are equivalent:

- (1) The probability measure is infinitely divisible.
- (2) For all $n \in \mathbb{N}$ the function $t \mapsto e^{\frac{1}{n}\psi(t)}$ is positive definite.
- (3) The cumulant function is conditionally positive definite.
- (4) The cumulant function has the following form

$$(1) \quad \psi(t) = i\langle t, \mu \rangle - \frac{1}{2}\langle \Gamma t, t \rangle + \int (e^{i\langle t, y \rangle} - 1 - i\langle t, h(y) \rangle) \nu(dy),$$

where

- (a) $h: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded, continuous, antisymmetric $h(-t) = -h(t)$, equal to the identity $h(t) = t$ in a neighborhood of θ^1 ;
- (b) $a, \mu \in \mathbb{R}$;
- (c) Γ is a symmetric and positive definite $n \times n$ matrix;

¹Other choices are possible, cf [2], [3].

(d) ν is a positive measure on \mathbb{R}^n such that $\nu(\{0\}) = 0$, for all $a > 0$, $\nu\{t \mid \|t\| > a\} < \infty$, and $\int \|y\|^2 \nu(dy) < \infty$.

Given h , the decomposition is unique and it is called Lévy-Khinchin formula. The triple (μ, Γ, ν) is called the Lévy triple and ν is called the Lévy measure.

Elements of the proof.

(1 \Rightarrow 4) As ϕ is infinitely divisible, there exists characteristic functions ϕ_n such that $\phi(t) = (\phi_n(t))^n$, $n = 1, 2, \dots$. From the existence and uniqueness of the cumulant function, $\phi = e^\psi$, as $\phi(t) = (\phi_n(t))^n = (e^{\phi_n})^n = e^{n\psi_n(t)}$, we obtain $\psi_n = \frac{1}{n}\psi$. The function $t \mapsto e^{n(\phi_n(t)-1)}$ is the characteristic function of the compound Poisson $\text{CP}(n, \mu_n)$, $\check{\mu}_n = \phi_n$, that is the characteristic function of $\sum_{k=0}^{N_n} X_k$, with $N_n \sim \text{Poi}(n)$, $X_0 = 0$, X_1, X_2, \dots IID μ_n , N_n, X_1, X_2, \dots independent. As

$$e^{n(\phi_n(t)-1)} = \exp\left(n(e^{\frac{1}{n}\psi(t)} - 1)\right) = \exp\left(\frac{e^{\frac{1}{n}\psi(t)} - 1}{1/n}\right) \rightarrow e^{\psi(t)} = \phi(t), \quad n \rightarrow \infty,$$

we have obtained that every infinitely divisible distribution is the limit of compound Poisson distribution, which are themselves infinitely divisible.

The cumulant function of the compound Poisson is

$$\psi_n^{(c)}(t) = \int (e^{ity} - 1) \nu_n(dy) = i\langle t, \mu_n \rangle + \int (e^{ity} - 1 - i\langle t, h(y) \rangle) \nu_n(dy),$$

with $\nu_n(dy) = n\mu_n(dy)$ and $\mu_n = \int h(y) \nu_n(dy)$. Note that the Gaussian term is missing. It is possibly produced in the limit as $n \rightarrow \infty$.

(4 \Rightarrow 3) The function $g(t) = i\langle t, \mu \rangle - \frac{1}{2}\langle \Gamma t, t \rangle$ is such that for all $\alpha > 0$ the function $f(t) = e^{\alpha g(t)}$ is the characteristic function of $N(\alpha\mu, \alpha\Gamma)$, hence positive definite. It follows that $(e^{\alpha g(t)} - 1)/\alpha$ is conditionally positive definite and so is $g(t) = \lim_{\alpha \downarrow 0} (e^{\alpha g(t)} - 1)/\alpha$. A direct computation shows that the integral part of (1) is conditionally positive definite.

(3 \Rightarrow 2) The function $f(t) = \frac{1}{n}\psi(t)$ is conditionally positive definite for all n and $f(0) = 0$.

Let us show that for all $t_1, \dots, t_m \in \mathbb{R}$ the matrix $[f(t_i - t_j) - f(t_i) - \overline{f(t_j)}]_{i,j=1}^m = [f(t_i - t_j) - f(t_i) - f(-t_j)]_{i,j=1}^m$ is positive definite. Chose $c_1, \dots, c_m \in \mathbb{C}$, $t_{m+1} = 0$, and c_{m+1} such that $\sum_{i=1}^{m+1} c_i = 0$. We have

$$\begin{aligned} \sum_{i,j=1}^m c_i \overline{c_j} (f(t_i - t_j) - f(t_i) - \overline{f(t_j)}) &= \\ &= \sum_{i,j=1}^m c_i \overline{c_j} f(t_i - t_j) - \sum_{i,j=1}^m c_i \overline{c_j} f(t_i) - \sum_{i,j=1}^m c_i \overline{c_j} f(-t_j) = \\ &= \sum_{i,j=1}^m c_i \overline{c_j} f(t_i - t_j) + \overline{c_{m+1}} \sum_{i=1}^m c_i f(t_i) + c_{m+1} \sum_{j=1}^m \overline{c_j} f(-t_j) = \\ &= \sum_{i,j=1}^{m+1} c_i \overline{c_j} f(t_i - t_j) \geq 0. \end{aligned}$$

The matrix $[e^{a_{ij}}]$ is positive definite if the matrix $[a_{ij}]$ is positive definite, then the matrix $[e^{f(t_i - t_j) - f(t_i) - \overline{f(t_j)}}]_{i,j=1}^m$ is positive definite and also

$$\sum_{i,j=1}^m c_i \overline{c_j} e^{f(t_i - t_j)} = \sum_{i,j=1}^m (c_i e^{f(t_i)}) \overline{c_j e^{f(t_j)}} e^{f(t_i - t_j) - f(t_i) - \overline{f(t_j)}} \geq 0$$

which shows that the function $t \mapsto e^f(t) = e^{\frac{1}{n}\psi(t)}$ is positive definite.

(2 \Rightarrow 1) The characteristic function $\phi(t) = e^{\psi(t)}$ is infinitely divisible because for all n
 $\phi(t) = (e^{\frac{1}{n}\psi(t)})^n$.

Uniqueness Assume there are two representations, so that

$$i\langle t, \mu \rangle_1 - \frac{1}{2}\langle \Gamma_1 t, t \rangle + \int (e^{i\langle t, y \rangle} - 1 - i\langle t, h(y) \rangle) \nu_1(dy) =$$

$$i\langle t, \mu \rangle_2 - \frac{1}{2}\langle \Gamma_2 t, t \rangle + \int (e^{i\langle t, y \rangle} - 1 - i\langle t, h(y) \rangle) \nu_2(dy),$$

hence

$$f(t) = \int (e^{i\langle t, y \rangle} - 1 - i\langle t, h(y) \rangle) \nu_1(dy) - \int (e^{i\langle t, y \rangle} - 1 - i\langle t, h(y) \rangle) \nu_2(dy),$$

where $f(t)$ is a polynomial of degree at most 2. The convolution of of this equality with $\chi = (\delta_0 - \delta_x)^{*3}$ gives

$$\int \hat{\chi}(y)e^{i\langle x, y \rangle} \nu_1(dy) = \int \hat{\chi}(y)e^{i\langle x, y \rangle} \nu_2(dy),$$

that is $(1 - e^{i\langle x, y \rangle})^3 \nu_1(dy) = (1 - e^{i\langle x, y \rangle})^3 \nu_2(dy)$, which implies $\nu_1 = \nu_2$. If the integrals are equal, then the first parts are equal. □

REFERENCES

1. Jean Jacod and Philip Protter, *Probability essentials*, second ed., Universitext, Springer-Verlag, Berlin, 2003. MR MR1956867 (2003m:60002)
2. Zoltán Sasvári, *Multivariate characteristic and correlation functions*, De Gruyter Studies in Mathematics, vol. 50, Walter de Gruyter & Co., Berlin, 2013. MR 3059796
3. Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 2013, Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR 3185174
4. David Williams, *Probability with martingales*, Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.

E-mail address: giovanni.pistone@carloalberto.org