LÉVY PROCESSES, STABLE PROCESSES, AND SUBORDINATORS

STEVEN P. LALLEY

1. DEFINITIONS AND EXAMPLES

Definition 1.1. A continuous–time process $\{X_t = X(t)\}_{t\geq 0}$ with values in $\mathbb{R}^d$ (or, more generally, in an abelian topological group G) is called a *Lévy process* if (1) its sample paths are right-continuous and have left limits at every time point t, and (2) it has stationary, independent increments, that is:

(a) For all $0 = t_0 < t_1 < \cdots < t_k$, the increments $X(t_i) - X(t_{i-1})$ are independent.
(b) For all $0 \leq s \leq t$ the random variables $X(t) - X(s)$ and $X(t-s) - X(0)$ have the same distribution.

The default initial condition is $X_0 = 0$. A *subordinator* is a real-valued Lévy process with nondecreasing sample paths. A *stable process* is a real-valued Lévy process $\{X_t\}_{t\geq 0}$ with initial value $X_0 = 0$ that satisfies the self-similarity property

$$X_t / t^{1/\alpha} \stackrel{\mathscr{D}}{=} X_1 \quad \forall\, t > 0. \tag{1.1}$$

The parameter α is called the *exponent* of the process.

Example 1.1. The most fundamental Lévy processes are the Wiener process and the Poisson process. The Poisson process is a subordinator, but is not stable; the Wiener process is stable, with exponent $\alpha = 2$. Any linear combination of independent Lévy processes is again a Lévy process, so, for instance, if the Wiener process W_t and the Poisson process N_t are independent then $W_t - N_t$ is a Lévy process. More important, linear combinations of independent Poisson processes are Lévy processes: these are special cases of what are called *compound Poisson processes*: see sec. 5 below for more. Similarly, if X_t and Y_t are independent Lévy processes, then the vector-valued process (X_t, Y_t) is a Lévy process.

Example 1.2. Let $\{W_t\}_{t\geq 0}$ be a standard Wiener process, and let $\tau(a)$ be the first passage time to the level $a > 0$, that is,

$$\tau(a) := \inf\{t : W_t > a\} \tag{1.2}$$

Then the process $\{\tau(a)\}_{a\geq 0}$ is a *stable subordinator* with exponent $\alpha = 1/2$. This follows from the strong Markov property of the Wiener process, which implies that the process $\tau(a)$ has stationary, independent increments, and the Brownian scaling property. (See sec. 3 below for a discussion of the strong Markov property.) The fact that $\tau(s)$ is right-continuous with left limits follows from the continuity of the Wiener path.

Date: April 12 2007.

Example 1.3. A *$d-$dimensional Wiener process* is an $\mathbb{R}^d-$valued process

$$W(t) = (W_1(t), W_2(t), \dots, W_d(t))$$

whose component $W_i(t)$ are independent standard one-dimensional Wiener processes. Let $W(t) = (X(t), Y(t))$ be a two-dimensional Wiener process, and let $\tau(s)$ be the first-passage process for the first coordinate, that is, $\tau(s) = \inf\{t : X(t) > s\}$. Then the process $\{C(s)\}_{s\geq 0}$ defined by

$$C(s) := Y(\tau(s)) \tag{1.3}$$

is a stable process with exponent 1; it is called the *Cauchy process*, because its increments have Cauchy distributions.

Exercise 1.1. (A) Prove that the process $C(s)$ is stable with exponent 1, using the strong Markov property of the two-dimensional Wiener process and the Brownian scaling property. (B) Check that

$$\exp\{i\theta Y_t - |\theta| X_t\}$$

is a continuous (complex-valued) martingale that remains bounded up to time $\tau(s)$. Then use the Optional Sampling theorem for bounded continuous martingales (essentially, the third Wald identity) to show that

$$E\exp\{i\theta Y(\tau(s))\} = \exp\{-|\theta| s\}. \tag{1.4}$$

This implies that the distribution of $C(s)$ is the *Cauchy distribution.*

Exercise 1.2. Let $\varphi(\lambda)$ be a nondecreasing (or alternatively a continuous) real-valued function of $\lambda \geq 0$ that satisfies $\varphi(0) = 1$ and the functional equation

$$\varphi(\lambda) = \varphi(m^{-1/\alpha})^m \quad \forall\, m \in \mathbb{N} \text{ and } \forall \lambda > 0.$$

Prove that for some constant $\gamma \geq 0$,

$$\varphi(\lambda) = \exp\{-\gamma\lambda^\alpha\}.$$

HINT: Start by making the substitution $f(r) = \varphi(r^{1/\alpha})$.

Exercise 1.3. (A) Show that if $X(t)$ is a stable subordinator with exponent α, then for some constant $\gamma \geq 0$,

$$Ee^{-\lambda X(t)} = \exp\{-\gamma t\lambda^\alpha\} \quad \forall\, t, \lambda > 0. \tag{1.5}$$

(B) Similarly, show that if $X(t)$ is a *symmetric* stable process with exponent α (here *symmetric* means that $X(t)$ has the same distribution as $-X(t)$), then for some constant $\gamma \geq 0$,

$$Ee^{i\theta X(t)} = \exp\{-\gamma t|\theta|^\alpha\}. \tag{1.6}$$

Exercise 1.4. (Bochner) Let $Y_t = Y(t)$ be a stable subordinator of exponent α, and let $W_t = W(t)$ be an independent standard Brownian motion. Use the strong Markov property (section 3 below) to show that

$$X(t) := W(Y(t)) \tag{1.7}$$

is a symmetric stable process of exponent 2α.

2. Infinitely Divisible Distributions and Characteristic Functions

Definition 2.1. A probability distribution F on $\mathbb{R}$ is said to be *infinitely divisible* if for every integer $n \geq 1$ there exist independent, identically distributed random variables $\{X_{n,i}\}_{1\leq i\leq n}$ whose sum has distribution F:

$$\sum_{i=1}^{n} X_{n,i} \stackrel{\mathscr{D}}{=} F. \tag{2.1}$$

Proposition 2.1. *If $\{X_t\}_{t\geq 0}$ is a Lévy process, then for each $t > 0$ the random variable X_t has an infinitely divisible distribution. Conversely, if F is an infinitely divisible distribution then there is a Lévy process such that X_1 has distribution F.*

Proof. The first statement is obvious, because by the definition of a Lévy process the increments $X((k+1)t/n) - X(kt/n)$ are independent and identically distributed. The converse is quite a bit trickier, and won't be needed for any other purpose later in the course, so I will omit it. □

Proposition 2.2. *Let $X_t = X(t)$ be a Lévy process, and for each $t \geq 0$ let $\varphi_t(\theta) = Ee^{i\theta X(t)}$ be the characteristic function of $X(t)$. Then there exists a continuous, complex-valued function $\psi(\theta)$ of $\theta \in \mathbb{R}$ such that for all $t \geq 0$ and all $\theta \in \mathbb{R}$,*

$$\varphi_t(\theta) = \exp\{t\psi(\theta)\}. \tag{2.2}$$

In particular, the function $\varphi_t(\theta)$ has no zeros.

Remark 1. In view of Proposition 2.1, it follows that every infinitely divisible characteristic function $\varphi(\theta)$ has the form $\varphi(\theta) = \exp\{\psi(\theta)\}$, and therefore has no zeros. In fact, though, the usual proof of the converse half of Proposition 2.1 proceeds by first showing that infinitely divisible characteristic functions have this form, and then using this to build the Lévy process. For the whole story, see the book *Probability* by Leo Breiman, chs. 9 and 14.

Proof of Proposition 2.2. The defining property of a Lévy process — that it has stationary, independent increments — implies that for each fixed θ, the characteristic function $\varphi_t(\theta)$ satisfies the multiplication rule

$$\varphi_{t+s}(\theta) = \varphi_t(\theta)\varphi_s(\theta).$$

Since a Lévy process has right-continuous sample paths, for each fixed argument θ the function $t \mapsto \varphi_t(\theta)$ is right-continuous, and in particular, since $\varphi_0(\theta) = 1$,

$$\lim_{t\to 0+} \varphi_t(\theta) = 1.$$

But this and the multiplication rule imply that the mapping $t \mapsto \varphi_t(\theta)$ must also be *left-continuous*.

The only continuous functions that satisfy the multiplication rule are the exponential functions $e^{\alpha t}$ and the zero function.[1] That $\varphi_t(\theta)$ is not identically zero (as a function of t) follows because $t \mapsto \varphi_t(\theta)$ is continuous at $t = 0$, where $\varphi_0(\theta) = 1$. □

Proposition 2.3. *Let $X(t)$ be a subordinator, and for each $t \geq 0$ let $\varphi_t(\lambda) = E^{-\lambda X(t)}$ be the Laplace transform of $X(t)$. Then there exists a continuous, nondecreasing, nonnegative function $\psi(\lambda)$ such that for all $t \geq 0$ and all $\lambda \geq 0$,*

$$\varphi_t(\lambda) = \exp\{-t\psi(\lambda)\}. \tag{2.3}$$

The proof is virtually the same as the proof of Proposition 2.2. The function $\psi(\lambda)$ associated with a subordinator has an interesting probabilistic interpretation, which will become at least partly clear in the discussion of Poisson point processes in sec. 5 below.

Proposition 2.2 implies that if X_t is a Lévy process then the characteristic function $Ee^{i\theta X_t}$ is continuous in t. This in turn implies that $X_s \Rightarrow X_t$ as $s \to t$ from the *left.* (Right-continuity of sample paths implies the stronger assertion that $X_s \to X_t$ almost surely as $s \to t$ from the right.) The weak convergence can be strengthened:

Proposition 2.4. *If $\{X_s\}_{s\geq 0}$ is a Lévy process then for each $t \geq 0$, the sample path X_s is continuous at $s = t$.*

Remark 2. Take note of how the statement is quantified! For each t, there is a null set on which the path may fail to be continuous at t. Since there are *uncountably* many t, these null sets might add up to something substantial. And indeed for some Lévy processes — e.g., the Poisson process — they do.

Proof. For each real θ, the process $Z_\theta(s) := \exp\{i\theta X_s - s\psi(\theta)\}$ is a martingale in s (relative to the natural filtration — see Definition 3.1 below). This is a routine consequence of the stationary, independent increments property of a Lévy process (exercise). Since $Z_\theta(s)$ is bounded for $s \in [0, t]$, it follows from the martingale convergence theorem that

$$\lim_{s\to t-} Z_\theta(s) = Z_\theta(t) \quad \text{almost surely.}$$

This implies that $e^{i\theta X_s} \to e^{i\theta X_t}$ almost surely as $s \to t-$, for every fixed θ, and therefore almost surely for every rational θ. Therefore, $X_s \to X_t$. (Explanation: If not, there would have to be a jump of size $2\pi k/\theta$ for some integer k, for every rational θ. But this rules out the possibility of a jump.) □

3. Strong Markov Property

Definition 3.1. Let $\{X(t)\}_{t\geq 0}$ be a Lévy process. The *natural filtration* associated to the process is the filtration $\mathscr{F}_t^X := \sigma(X_s)_{s\leq t}$, that is, each σ−algebra $\mathscr{F}_t^X$ is the smallest σ−algebra with respect to which all of the random variables $X(s)$, for $s \leq t$, are measurable. An *admissible filtration* $\{\mathscr{F}_t\}_{t\geq 0}$ for the process $\{X(t)\}_{t\geq 0}$ is a filtration such that

(a) $\mathscr{F}_t^X \subseteq \mathscr{F}_t$ for each t, and

[1] Exercise! In fact, the only *Lebesgue measurable* functions $f(t)$ that satisfy the *addition* rule $f(t+s) = f(s) + f(t)$ are the linear functions $f(t) = a + bt$: this is a considerably harder exercise.

(b) each $\sigma-$algebra $\mathscr{F}_t$ is independent of the $\sigma-$algebra

$$\mathscr{G}_t^X := \sigma(X(t+s) - X(t))_{s\geq 0}. \tag{3.1}$$

For any filtration $\mathbb{F} = \{\mathscr{F}_t\}_{t\geq 0}$, set

$$\mathscr{F}_\infty = \sigma(\cup_{t\geq 0}\mathscr{F}_t), \tag{3.2}$$

that is, the smallest $\sigma-$algebra containing all events measurable relative to some $\mathscr{F}_t$.

The reason for introducing the notion of an admissible filtration is that it allows inclusion of events determined by other independent processes. For instance, if $X(t)$ and $Y(t)$ are independent Wiener processes, as in Example 1.3 above, then the natural filtration $\mathscr{F}_t^{(X,Y)}$ for the vector-valued process (X_t, Y_t) will be admissible for each of the coordinate processes.

Definition 3.2. Let $\mathbb{F} = \{\mathscr{F}_t\}_{t\geq 0}$ be a filtration and let τ be a stopping time relative to $\mathbb{F}$. (Recall that a nonnegative random variable τ is a stopping time relative to a filtration $\mathbb{F}$ if for every $t \geq 0$ the event $\{\tau \leq t\} \in \mathscr{F}_t$.) The *stopping field* $\mathscr{F}_\tau$ induced by τ is the collection of all events A such that for each $t \geq 0$,

$$A \cap \{\tau \leq t\} \in \mathscr{F}_t. \tag{3.3}$$

Theorem 1. *(Strong Markov Property I) Let $\{X(t)\}_{t\geq 0}$ be a Lévy process and $\mathbb{F} = \{\mathscr{F}_t\}$ and admissible filtration. Suppose that τ is a stopping time relative to $\mathbb{F}$. Define the post-τ process*

$$Y(t) := X(\tau + t) - X(\tau). \tag{3.4}$$

Then the process $\{Y(t)\}_{t\geq 0}$ is independent of $\mathscr{F}_\tau$ and identical in law to $\{X(t)\}_{t\geq 0}$. In detail, for every event $A \in \mathscr{F}_\tau$, all t_i, and all Borel sets B_i,

$$\begin{aligned} P(A \cap \{Y(t_i) \in B_i \ \forall\, i \leq k\}) &= P(A)P\{X(t_i) \in B_i \ \forall\, i \leq k\} \\ &= P(A)P\{Y(t_i) \in B_i \ \forall\, i \leq k\} \end{aligned} \tag{3.5}$$

Theorem 2. *(Strong Markov Property II) Suppose in addition to the hypotheses of Theorem 1 that $\{X_t^*\}_{t\geq 0}$ is identical in law to $\{X(t)\}_{t\geq 0}$ and is independent of the stopping field $\mathscr{F}_\tau$. Define the* spliced process

$$\begin{aligned} \tilde{X}_t &:= X_t \quad \text{if } t < \tau \quad \text{and} \\ \tilde{X}_t &:= X_\tau + X^*_{t-\tau} \quad \text{if } t \geq \tau. \end{aligned} \tag{3.6}$$

Then the process $\tilde{X}_t$ is also a version of (identical in law to) X_t.

Remark 3. The identity (3.5) is equivalent to the following property: for every $k \geq 1$, every bounded continuous function $f : \mathbb{R}^k \to \mathbb{R}$, and every event $A \in \mathscr{F}_\tau$,

$$\begin{aligned} E\mathbf{1}_A f(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) &= P(A)Ef(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) \\ &= P(A)Ef(X_{t_1}, X_{t_2}, \dots, X_{t_k}) \end{aligned} \tag{3.7}$$

This equivalence follows by a standard approximation argument in measure theory (that in essence asserts that indicator functions of Borel events can be arbitrarily well approximated in L^1 by bounded continuous functions). See Proposition **??** in section **??** below. Moreover, the identity (3.5) implies that the $\sigma-$algebras $\mathscr{F}_\tau$ and $\mathscr{F}^Y_\infty$ are independent. This follows from the $\pi-\lambda$ theorem, because (a) the collection of events $\{Y(t_i)\in B_i \ \forall\, i\le k\}$ is a $\pi-$system; and (b) for any event $A\in\mathscr{F}_\tau$, the collection of events B for which

$$P(A\cap B)=P(A)P(B)$$

is a $\lambda-$system.

Proof of Theorem 1. The strategy is to first prove the result for *discrete* stopping times, and then to deduce the general case by an approximation argument. First consider the special case where the stopping time takes values in a countable, discrete set $S=\{s_i\}_{i\in\mathbb{N}}$, with $s_i<s_{i+1}$. (Here *discrete* means that the set S has no limit points in $\mathbb{R}$, that is, $\lim_{n\to\infty}s_n=\infty$.) If this is the case then for each i the event $\{\tau=s_i\}\in\mathscr{F}_{s_i}$, and to check (3.5) it suffices to consider events $A=G\cap\{\tau=s_i\}$ where $G\in\mathscr{F}_{s_i}$. For such events A, the equality (3.5) follows easily from the hypothesis that the filtration is admissible and the fact that a Lévy process has stationary, independent increments (check this!). Thus, the Strong Markov Property I is valid for discrete stopping times, and it follows by the note above that identity (3.7) holds for discrete stopping times.

Now let τ be an arbitrary stopping time. Approximate τ from above by discrete stopping times τ_n, for instance by setting $\tau_n=$ the smallest dyadic rational $k/2^n$ larger than τ. Each such τ_n is a stopping time (check this), and

$$\lim_{n\to\infty}\downarrow\tau_n=\tau.$$

Moreover, if $A\in\mathscr{F}_\tau$ then $A\in\mathscr{F}_{\tau_n}$ for each n. (See Exercise 3.1 below.) Consequently, by the previous paragraph, the equality (3.7) holds when $Y(t_i)$ is replaced by $Y_n(t_i)$, where $Y_n(t):=X(\tau_n+t)-X(\tau_n)$. But by the right-continuity of sample paths,

$$\lim_{n\to\infty}Y_n(t)=Y(t)\quad a.s.,$$

and so if $f:\mathbb{R}^k\to\mathbb{R}$ is a bounded continuous function then

$$\lim_{n\to\infty}f(Y_n(t_1),Y_n(t_2),\dots,Y_n(t_k))=f(Y(t_1),Y(t_2),\dots,Y(t_k))$$

Therefore, by the dominated convergence theorem, (3.7) follows from the fact that it holds when Y is replaced by Y_n. □

Exercise 3.1. Prove that if τ and σ are stopping times (relative to the same filtration $\mathbb{F}$) such that $\sigma\le\tau$, then $\mathscr{F}_\sigma\subset\mathscr{F}_\tau$.

Proof of Theorem 2. This is more subtle than Theorem 1, because a straightforward attempt to deduce the general case from the special case of discrete stopping times doesn't work. The difficulty is that if one approximates τ from above by τ_n, as in the proof of Theorem 1, then the process X^*_t may not be independent of the stopping fields $\mathscr{F}_{\tau_n}$ (these *decrease* to $\mathscr{F}_\tau$ as $n\to\infty$).

So we'll take a different tack, based on the optional sampling formula for martingales. To prove the theorem, we must show that the finite-dimensional distributions of the spliced process $\tilde{X}_t$ are the same as those of X_t. For this, it suffices to show that the joint characteristic functions agree, that is, for all choices of $\theta_j \in \mathbb{R}$ and all times t_j,

$$E\exp\left\{i\sum_{j=1}^{k}\theta_j(X_{t_{j+1}}-X_{t_j})\right\}=E\exp\left\{i\sum_{j=1}^{k}\theta_j(\tilde{X}_{t_{j+1}}-\tilde{X}_{t_j})\right\}.$$

To simplify the exposition, I'll consider only the case $k=1$; the general case can be done in the same manner. So: the objective is to show that for each $\theta\in\mathbb{R}$ and each $t>0$,

$$Ee^{i\theta X_t}=Ee^{i\theta\tilde{X}_t}.$$

In proving this, we may assume that the stopping time τ satisfies $\tau\le t$, because if it doesn't we can replace it by the smaller stopping time $\tau\wedge t$. (The stopping field $\mathscr{F}_{\tau\wedge t}$ is contained in $\mathscr{F}_\tau$, so if the process X^*_s is independent of $\mathscr{F}_\tau$ then it is also independent of $\mathscr{F}_{\tau\wedge t}$.) The stationary, independent increments property implies that the process $Z_\theta(s):=\exp\{i\theta X_s-s\psi(\theta)\}$ is a martingale relative to the filtration $\mathscr{F}_s$, and for each fixed θ this process is bounded for $s\le t$. Consequently, the optional sampling formula gives

$$Ee^{i\theta X_\tau-\tau\psi(\theta)}=1.$$

Thus, to complete the proof of (**??**) it suffices to show that

$$E(\exp\{i\theta X^*(t-\tau)-(t-\tau)\psi(\theta)\}\,|\,\mathscr{F}_\tau)=1. \tag{3.8}$$

The proof of (3.8) will also turn on the optional sampling formula, together with the hypothesis that the process X^*_s is independent of $\mathscr{F}_\tau$. Let $\mathscr{F}^*_s=\sigma(X^*_r)_{r\le s}$ be the natural filtration of X^*, and set

$$\mathscr{G}_s=\sigma(\mathscr{F}^*_s\cup\mathscr{F}_\tau).$$

This is an *admissible* filtration for the process X^*, because X^* is independent of $\mathscr{F}_\tau$. The random variable $t-\tau$ is measurable relative to $\mathscr{G}_0=\mathscr{F}_\tau$, and so it is (trivially) a stopping time relative to the filtration $\mathscr{G}_s$. Because the filtration $\mathscr{G}_s$ is admissible for the process X^*_s, the stationary, independent increments property implies that the process

$$Z^*_\theta(s):=\exp\{i\theta X^*_s-s\psi(\theta)\}$$

is a martingale relative to the filtration $\mathscr{G}_s$, and remains bounded up to time t. Thus, the optional sampling formula (applied with the stopping time $t-\tau$) implies (3.8). □

4. BLUMENTHAL ZERO-ONE LAW

Theorem 3. *(Blumenthal Zero-One Law) Let $\{X(t)\}_{t\ge0}$ be a Lévy process and $\{\mathscr{F}^X_t\}_{t\ge0}$ its natural filtration. Define the $\sigma-$algebra*

$$\mathscr{F}^X_{0+}:=\bigcap_{t>0}\mathscr{F}^X_t. \tag{4.1}$$

Then $\mathscr{F}^X_{0+}$ is a zero-one field, that is, every event $A\in\mathscr{F}^X_{0+}$ has probability zero or one.

Proof. For notational ease, I'll drop the superscript X from the $\sigma-$algebras. Set $\mathscr{F}_\infty = \sigma(\cup_{t\geq 0}\mathscr{F}_t)$. I claim that for any event $A\in\mathscr{F}_\infty$,

$$E(\mathbf{1}_A\,|\,\mathscr{F}_{0+}) = P(A) \text{ a.s.} \tag{4.2}$$

This will imply the theorem, because for any event $A\in\mathscr{F}_{0+}$ it is also the case that

$$E(\mathbf{1}_A\,|\,\mathscr{F}_{0+}) = \mathbf{1}_A \text{ a.s.},$$

by the filtering property of conditional expectation. To prove (4.2), it is enough to consider cylinder events, that is, events of the form

$$A = \bigcap_{i=1}^{k}\{X(t_i)\in B_i\} \quad \text{where } B_i\in\mathscr{B}_1.$$

Let $f:\mathbb{R}^k\to\mathbb{R}$ be a bounded, continuous function, and set

$$\xi := f(X(t_1), X(t_2), \dots, X(t_k)) \quad \text{and}$$
$$\xi_n := f(X(t_1)-X(1/n), X(t_2)-X(1/n), \dots, X(t_k)-X(1/n)).$$

To prove the identity (4.2) for cylinder events, it is enough (by Exercise 4.1 below) to prove that for any bounded continuous $f:\mathbb{R}^k\to\mathbb{R}$,

$$E(\xi\,|\,\mathscr{F}_{0+}) = E\xi. \tag{4.3}$$

Since the process $X(t)$ has right-continuous paths, $\xi_n\to\xi$ a.s., and the convergence is bounded, because f was assumed to be bounded. Therefore, by the usual dominated convergence theorem and the DCT for conditional expectations,

$$\lim_{n\to\infty} E\xi_n = E\xi \quad \text{and}$$
$$\lim_{n\to\infty} E(\xi_n\,|\,\mathscr{F}_{0+}) = E(\xi\,|\,\mathscr{F}_{0+}) \quad \text{a.s.}$$

Therefore, to prove (4.3) it suffices to show that for each n,

$$E(\xi_n\,|\,\mathscr{F}_{0+}) = E\xi_n \quad \text{a.s.}$$

But this follows immediately from the independent increments property, because ξ_n is a function of the increments after time $1/n$, and so is independent of $\mathscr{F}_{1/n}$ – and hence also independent of $\mathscr{F}_{0+}$, since this is contained in $\mathscr{F}_{1/n}$.

□

Exercise 4.1. Complete the proof above by showing that the identity (4.3) implies the identity 4.2 for all cylinder events.

Corollary 4.1. *If $X(t)$ is a subordinator with continuous sample paths then there is a constant $C\geq 0$ such that $X(t) = Ct$ a.s.*

Proof. According to a basic theorem of real analysis (see, for instance, H. ROYDEN, *Real Analysis*, chapter 5), every nondecreasing function is differentiable a.e. Hence, in particular, $X(t)$ is differentiable in t a.e., with probability one. It follows that $X(t)$ is differentiable at $t=0$ almost surely, and that the derivative is finite a.s. (See the technical note following the proof for further elaboration.) But the derivative

$$X'(0) := \lim_{\varepsilon \to 0+} \frac{X(\varepsilon)}{\varepsilon}$$

depends only on the values of the process $X(t)$ for t in arbitrarily small neighborhoods of 0, and so it must be $\mathscr{F}_{0+}^X$−measurable. Since $\mathscr{F}_{0+}^X$ is a zero-one field, by Theorem 3, it follows that the random variable $X'(0)$ is constant a.s: thus, $X'(0) = C$ for some constant $C \geq 0$.

Existence of a finite derivative C may be restated in equivalent geometric terms as follows: for any $\delta > 0$, the graph of $X(t)$ must lie entirely in the cone bounded by the lines of slope $C \pm \delta$ through the origin, at least for small $t > 0$. Thus, if we define

$$\begin{aligned} T = T_1 : &= \min\{t > 0 : X(t) \geq (C + \delta)t\} \\ &= \infty \quad \text{if there is no such } t \end{aligned}$$

then $T > 0$ (but possibly infinite) with probability one. Note that T is a stopping time, by path continuity. Moreover, since $X(t)$ is continuous, by hypothesis, $X(T) = (C + \delta)T$ on the event $T < \infty$ and $X(t) \leq Ct$ for all $t \leq T$. Now define inductively

$$\begin{aligned} T_{k+1} &= \min\{t > 0 : X(t + T_k) - X(T_x) \geq (C + \delta)t\} \\ &= \infty \quad \text{if } T_k = \infty \text{ or if no such } t \text{ exists.;} \end{aligned}$$

observe that $X(T_k) = (C+\delta)T_k$ on the event $T_k < \infty$, for each k, and so $X(t) \leq (C+\delta)t$ for all $t \leq T_k$. By the Strong Markov Property, the increments $T_{k+1} - T_k$ are i.i.d. and strictly positive. Hence, by the SLLN, $T_k \to \infty$ almost surely as $k \to \infty$, so $X(t) \leq (C + \delta)t$ for all $t < \infty$. The same argument shows that $X(t) \geq (C - \delta)t$ for all $t \geq 0$. Since $\delta > 0$ is arbitrary, it follows that $X(t) = Ct$ for all t, with probability one. □

Remark 4. The random process $X(t)$ is actually a function $X(t, \omega)$ of two arguments $t \in [0, \infty)$ and $\omega \in \Omega$. Consider only those $t \in [0, 1]$; then $X(t, \omega)$ can be viewed as a single random variable on the product space $[0, 1] \times \Omega$, endowed with the product probability measure Lebesgue×P. The set of all pairs (t, ω) for which the derivative $X'(t)$ exists is product measurable (why?). By the differentiation theorem quoted above, for P−a.s. ω the derivative$X'(t)$ exists for a.e. t; hence, by Fubini's theorem, for a.e. t the derivative $X'(t)$ exists *at* t with P−probability one. But by the stationary independent increments property, the events

$$G_t := \{\omega : X(\cdot, \omega) \text{ differentiable at } t\}$$

all have the same probability. Therefore, $P(G_t) = 1$ for every t, and in particular, for $t = 0$.

5. Lévy Processes from Poisson Point Processes

5.1. **Compound Poisson Processes.**

Definition 5.1. A *compound Poisson process* $X(t)$ is a vector-valued process of the form

$$X(t) = \sum_{i=1}^{N(t)} Y_i, \tag{5.1}$$

where $N(t)$ is a Poisson counting process of rate $\lambda > 0$ and $\{Y_i\}_{i\geq 1}$ are independent, identically distributed random vectors independent of the Poisson process $N(t)$. The distribution F of the increments Y_i is called the *compounding distribution.*

Every compound Poisson process is a Lévy process. However, since compound Poisson processes have sample paths that are step functions, not all Lévy processes are compound Poisson processes. But every Lévy process can be *approximated* arbitrarily closely by compound Poisson processes plus linear drift. We will show in detail below how this approximation works for subordinators.

There are simple formulas for the characteristic function and (in case the increments Y_i are nonnegative) the Laplace transform of the random variables $X(t)$ that constitute a compound Poisson process. To obtain these, just condition on the number $N(t)$ of jumps in the underlying Poisson process, and use the fact that the conditional distribution of $X(t)$ given $N(t) = n$ is the $n-$fold convolution of the compounding distribution F. The results are:

$$Ee^{i\langle\theta,X(t)\rangle} = \exp\{\lambda t\varphi_F(\theta) - \lambda t\} \quad \text{where} \quad \varphi_F(\theta) = Ee^{i\langle\theta,Y_1\rangle} \tag{5.2}$$

and similarly, if the jumps Y_i are nonnegative,

$$Ee^{-\beta X(t)} = \exp\{\lambda t\psi_F(\beta) - \lambda t\} \quad \text{where} \quad \psi_F(\beta) = Ee^{-\beta Y_1}. \tag{5.3}$$

5.2. **Poisson point processes.** There is another way of looking at the sums (5.1) that define a compound Poisson process. Place a point in the first quadrant of the plane at each point (T_i, Y_i), where T_i is the time of the ith jump of the driving Poisson process $N(t)$, and call the resulting random collection of points $\mathscr{P}$. Then equation (5.1) is equivalent to

$$X(t) = \sum_{(T_i,Y_i)\in\mathscr{P}} Y_i \mathbf{1}\{T_i \leq t\} \tag{5.4}$$

This equation says that $X(t)$ evolves in time as follows: slide a vertical line to the right at speed one, and each time it hits a point (T_i, Y_i) add Y_i. The random collection of points (T_i, Y_i) constitutes a *Poisson point process* (also called a *Poisson random measure*) in the plane.

Definition 5.2. Let $(\mathscr{X}, \mathscr{B}, \mu)$ be a measure space. A *Poisson point process* $\mathscr{P}$ on $(\mathscr{X}, \mathscr{B})$ with ($\sigma-$finite) intensity measure μ is a collection $\{N(B)\}_{B\in\mathscr{B}}$ of extended nonnegative integer-valued random variables such that

(A) If $\mu(B) = \infty$ then $N(B) = \infty$ a.s.
(B) If $\mu(B) < \infty$ then $N(B) \sim$Poisson-($\mu(B)$).
(C) If $\{B_i\}_{i\in\mathbb{N}}$ are pairwise disjoint, then
 (C1) the random variables $N(B_i)$ are independent; and
 (C2) $N(\cup_i B_i) = \sum_i N(B_i)$.

Note: Condition (C2) asserts that the assignment $B \mapsto N(B) = N(B,\omega)$ is countably additive, that is, for each fixed $\omega \in \Omega$ the set function $N(B,\omega)$ is a *measure.*

Definition 5.3. Let $(\mathscr{X}, \mathscr{B}, \mu)$ be a measure space. A *random measure* on $(\mathscr{X}, \mathscr{B})$ with *intensity* μ is a family of $[0,\infty]-$valued random variables $\{M(B)\}_{B\in\mathscr{B}}$ such that

(A) For each $\omega \in \Omega$ the set function $B \mapsto M(B,\omega)$ is a measure on $(\mathscr{X}, \mathscr{B})$, and
(B) $EM(B) = \mu(B)$ for every $B \in \mathscr{B}$.

Clearly, a Poisson point process is a random measure. Note that in general the random variables $M(B)$ need not be integer-valued; when they are, the random measure M is called a *point process*. In the case of a compound Poisson process, the relevant intensity measure on $\mathbb{R}^2$ is

$$\mu(dt, dy) = \lambda \, dt \, dF(y) \tag{5.5}$$

Note that to each strip $\Gamma_t := \{(s,y) : s \leq t\}$ the intensity measure assigns (finite) mass λt. In addition, property (c) implies that the assignment $B \mapsto N(B) = N(B,\omega)$ is a measure for almost every ω. (Exercise: Why?) Since $N(B)$ is just a count of the number of points in the Poisson point process inside B, integration against the random measure N amounts to nothing more than summing over the collection of points $\mathscr{P}$. In particular, (5.4) can be interpreted as an integral:

$$\boxed{X(t) = \iint y \mathbf{1}_{[0,t]}(s) N(ds, dy);} \tag{5.6}$$

in the case of a compound Poisson process, this integral always reduces to a finite sum. Now in principle there is no reason that we can't use equation (5.6) to define a stochastic process $\{X(t)\}$ for more general Poisson random measures N on the plane – the only difficulty is that, since the integral (5.6) is no longer necessarily a finite sum, it might not be well-defined.

5.3. **Subordinators and Poisson Point Processes.**

Proposition 5.1. *Let $N(B)$ be a Poisson point process on the first quadrant whose intensity measure $\mu(dt, dy)$ has the form*

$$\mu(dt, dy) = dt \, \nu(dy) \tag{5.7}$$

for some Borel measure $\nu(dy)$ on the positive halfline $\mathbb{R}_+$, called the Lévy measure. *Assume that the Lévy measure ν satisfies*

$$\int (y \wedge 1) \nu(dy) < \infty. \tag{5.8}$$

Then the (random) integral in (5.6) *is well-defined and finite almost surely for each $t \geq 0$, and the process $X(t)$ defined by* (5.6) *is a* subordinator. *Moreover, the Laplace transform of the random variable $X(t)$ is given by*

$$Ee^{-\beta X(t)} = \exp\left\{ t \int_0^\infty (e^{-\beta y} - 1) \nu(dy) \right\}. \tag{5.9}$$

Remark 5. Formula (5.9) agrees with equation (5.3) in the special case where $\nu(dy) = \lambda F(dy)$, so (5.9) holds for compound Poisson processes.

Before proceeding to the proof, let's consider an important example:

Corollary 5.2. *For each $\alpha \in (0,1)$ the measure*

$$\nu(dy) = \nu_\alpha(dy) = y^{-\alpha-1}\,dy \quad \text{on}\,(0,\infty). \tag{5.10}$$

is the Lévy measure of a stable subordinator of exponent α.

Proof. It is easily checked that the measure ν_α defined by (5.10) will satisfy the integrability hypothesis (5.8) if and only if $0 < \alpha < 1$. For all such α, Proposition 5.1 asserts that formula (5.6) defines a subordinator. I will show that this subordinator is a stable process of exponent α by using the formula (5.9) to evaluate its Laplace transform. (In particular, we will see that the Laplace transform coincides with (1.5) above).

According to formula (5.9),

$$\log E e^{-\beta X(t)} := t\psi(\beta) = t\int_0^\infty (e^{-\beta y} - 1)\frac{dy}{y^{\alpha+1}}.$$

Make the linear substitution $x = ry$ to check that the function ψ satisfies the functional equation $\psi(r\beta) = r^\alpha \psi(\beta)$ for all $r, \beta > 0$, and conclude from this by calculus that

$$\psi(\beta) = \beta^\alpha \psi(1) = -\beta^\alpha(\alpha^{-1}\Gamma(1-\alpha)).$$

Thus, with $\gamma = \alpha^{-1}\Gamma(1-\alpha)$,

$$E e^{-\beta X(t)} = \exp\{-\gamma t \beta^\alpha\}.$$

This implies that $X(t)$ satisfies the scaling law (1.1), and therefore is a stable subordinator of exponent α. □

Exercise 5.1. (A) Verify that the functional equation $\psi(r\beta) = r^\alpha\psi(\beta)$ implies that $\psi(\beta) = \beta^\alpha\psi(1)$. (B) Verify that $\gamma = \alpha^{-1}\Gamma(1-\alpha)$. (C) How did I guess that $y^{-\alpha-1}\,dy$ is the right Lévy measure to produce a stable subordinator? HINT: Use the scaling law (1.1).

Proof of Proposition 5.1. First, we must show that the integral in (5.6) is well-defined and finite a.s. Since the intensity measure $dt\,\nu(dy)$ is concentrated on the first quadrant, the points of the Poisson point process are all in the first quadrant, with probability one. Consequently, the integral in (5.6) is a sum whose terms are all nonnegative, and so it is well-defined, although possibly infinite. To show that it is in fact finite a.s., we break the sum into two parts:

$$X(t) = \iint_{y>1} y\mathbf{1}_{[0,t]}(s)N(ds,dy) + \iint_{0\le y\le 1} y\mathbf{1}_{[0,t]}N(ds,dy). \tag{5.11}$$

Consider the first integral: The number of points of the Poisson point process $\mathscr{P}$ in the region $[0,t]\times(1,\infty)$ is Poisson with mean $\int_{y>1}\nu(dy)$, and this mean is finite, by the hypothesis (5.8). Consequently the first integral in (5.11) has only finitely many terms, with probability one, and therefore is finite a.s. Now consider the second integral in (5.11). In general, this integral is a sum with possibly infinitely many terms; to show that it is finite a.s., it is enough to show that its expectation is finite. For this, observe that

$$\iint_{0\le y\le 1} y\mathbf{1}_{[0,t]}(s)N(ds,dy) \le \sum_{k=0}^\infty 2^{-k} N([0,t]\times[2^{-k-1},2^{-k}]).$$

But

$$2^{-k}EN([0,t]\times[2^{-k-1},2^{-k}])=t\int_{[2^{-k-1},2^{-k}]}2^{-k}\nu(dy)\le 2t\int_{[2^{-k-1},2^{-k}]}y\nu(dy),$$

and so

$$E\iint_{0\le y\le 1}y\mathbf{1}_{[0,t]}(s)N(ds,dy)\le 2t\int_{[0,1]}y\nu(dy)<\infty,$$

by hypothesis (5.8). This proves that the random integral in (5.6) is well-defined and finite a.s. Clearly the integral is nonnegative, and nondecreasing in t. Thus, to show that the process $X(t)$ is a subordinator it is enough to prove that $X(t)$ is right-continuous in t. This, however, follows directly from the dominated convergence theorem for the random measure $N(B)$. (Exercise: Fill in the details here.)

Finally, consider the Laplace transform $Ee^{-\beta X(t)}$. To see that it has the form (5.9), we use the fact noted earlier that (5.9) is true in the special case of a compound Poisson process, and obtain the general case by an approximation argument. The key is that the random variable $X(t)$ may be approximated from below as follows:

$$X(t)=\lim_{\varepsilon\to 0}X_\varepsilon(t)\quad\text{where}\quad X_\varepsilon(t)=\iint_{y\ge\varepsilon}y\mathbf{1}_{[0,t]}(s)N(ds,dy).$$

Note that $X_\varepsilon(t)$ is obtained from $X(t)$ by removing all of the jumps of size less than ε. This makes $X_\varepsilon(t)$ a compound Poisson process, with compounding distribution

$$F(dy)=\mathbf{1}_{[\varepsilon,\infty)}\nu(dy)/\nu([\varepsilon,\infty))$$

and $\lambda=\nu([\varepsilon,\infty))$. Therefore, the Laplace transform of $X_\varepsilon(t)$ is given by

$$Ee^{-\beta X_\varepsilon(t)}=\exp\left\{t\int_\varepsilon^\infty(e^{-\beta y}-1)\nu(dy)\right\}.$$

Since the random variables $X_\varepsilon(t)$ are nonnegative and nonincreasing in ε, it follows that

$$Ee^{-\beta X(t)}=\lim_{\varepsilon\to 0}Ee^{-\beta X_\varepsilon(t)},$$

and so the representation (5.9) follows by monotone convergence. □

5.4. **Symmetric Lévy Processes and PPPs.** In Proposition 5.1, the intensity measure $\mu(dt,dy)=dt\,\nu(dy)$ is concentrated entirely on the first quadrant, and so the jumps of the process $X(t)$ defined by (5.6) are all positive. It is also possible to build Lévy processes with both positive and negative jumps according to the recipe (5.6), but as in Proposition 5.1 some restrictions on the intensity measure of the Poisson point process are necessary to ensure that the infinite series produced by (5.6) will converge. Following is a condition on the intensity measure appropriate for the construction of *symmetric* Lévy processes, that is, Lévy processes $X(t)$ such that for each t the distribution of $X(t)$ is the same as that of $-X(t)$. For the complete story of which intensity measures can be used to build Lévy processes in general, see J. Bertoin, *Lévy Processes,* ch. 1.

Proposition 5.3. *Let $N(B)$ be a Poisson point process on the first quadrant whose intensity measure $\mu(dt,dy)$ has the form*

$$\mu(dt,dy)=dt\,\nu(dy) \tag{5.12}$$

for some Borel measure $\nu(dy)$ on the real line. Assume that $\nu(dy)$ is symmetric*, that is, that $\nu(dy)=\nu(-dy)$, and that*

$$\int (y^2\wedge 1)\nu(dy)<\infty. \tag{5.13}$$

Then the (random) integral in (5.6) *is well-defined and finite almost surely for each $t\geq 0$, and the process $X(t)$ defined by* (5.6) *is a symmetric Lévy process. Moreover, the characteristic function of $X(t)$ is given by*

$$Ee^{i\theta X(t)}=\exp\left\{t\int_{\mathbb{R}}(e^{i\theta y}-1)\nu(dy)\right\}. \tag{5.14}$$

Proof. Omitted. See Bertoin, *Lévy Processes*, ch. 1 for a more general theorem along these lines. □

Exercise 5.2. Consider the measure $\nu(dy)=y^{-\alpha-1}\,dy$ for $y\in\mathbb{R}$. For which values of α does this measure satisfy the hypothesis of Proposition 5.3? Show that for these values, the corresponding Lévy process is a symmetric stable process.

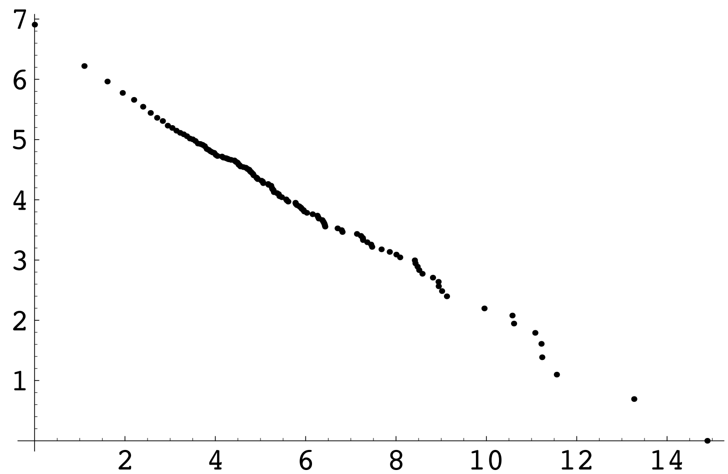
5.5. **Stable Processes and Power Laws.** The representation of stable subordinators and symmetric stable processes by Poisson point processes as given in Corollary 5.2 and Exercise 5.2 has some important implications about jump data for such processes. Suppose, for instance, that you were able to record the size of successive jumps in a stable-α subordinator. Of course, the representation of Corollary 5.2 implies that there are infinitely many jumps in any finite time interval; however, all but finitely many are of size less than (say) 1. So consider only the data on jumps of size at least 1. What would such data look like?

The Poisson point process representation implies that the number $M(y,t)$ of jumps of size at least y occurring before time t is Poisson with mean

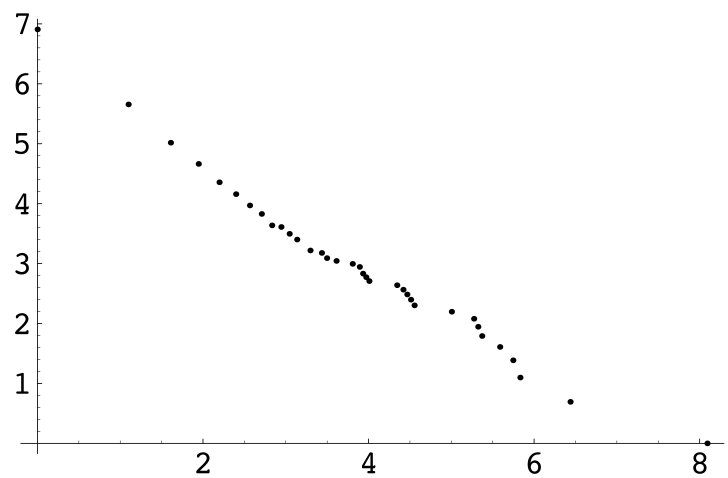
$$tG(y):=t\int_y^\infty u^{-\alpha-1}\,du=Cty^{-\alpha}.$$

This is what physicists call a *power law*, with exponent α. If you were to plot $\log G(y)$ against $\log y$ you would get a straight line with slop $-\alpha$. If you have enough data from the corresponding Poisson point process, and you make a similar log/log plot of the *empirical* counts, you will also see this straight line. Following is an example for the stable-1/2 subordinator. I ran 1000 simulations of a simple random walk excursion, that is, each run followed a simple random walk until the first visit to $+1$. I recorded the duration L of each excursion. (Note: Roughly half of the excursions had $L=1$, for obvious reasons; the longest excursion had $L=2,925,725$.) I sorted these in increasing order of

L, then, for each recorded L, counted $N(L)$ = number of excursions of duration at least L (thus, for example, $N(1) = 1,000$ and $N(3) = 503$). Here is a plot of $\log N(L)$ versus $\log L$:



This data clearly follows a line of slope $-1/2$. Next is a similar simulation for the Cauchy process. To (approximately) simulate the jumps of a Cauchy process, start with the jumps J of a stable-1/2 process, as constructed (approximately) above, and for each value J, run an independent simple random walk for J steps, and record the value S of the endpoint. I did this for the 1,000 jump values plotted in the first figure above. Here is a log/log plot of the resulting $|S|$−values:



This data follows a line of slope -1, just as theory predicts. If you are wondering why there seem to be only half as many points in this plot as in the previous plot, it is because I grouped the S− values by their *absolute* values, and there were only about half as many distinct $|S|$−values as S−values. In case you are interested, I have left the MATHEMATICA notebook on the web page.

5.6. **Representation Theorem for Subordinators.**

Theorem 4. *Every subordinator $Y(t)$ has the form $Y(t) = Ct + X(t)$, where $C \geq 0$ is a nonnegative constant and $X(t)$ is a subordinator of the form* (5.6)*, where $N(dt, dy)$ is a Poisson point process whose intensity measure $\mu(dt, dy) = dt\,\nu(dy)$ satisfies* (5.8).

There is a similar representation for the general one-dimensional Lévy process, called the *Lévy-Khintchine* representation. This asserts that in general a Lévy process is of the form

$$X(t) = Ct + \sigma W(t) + \iint y\,\mathbf{1}_{[0,t]}(s) N(ds, dy), \tag{5.15}$$

where N is a Poisson point process whose intensity measure satisfies the hypotheses of Proposition 5.3 and $W(t)$ is an independent Wiener process.

Proof. (Sketch) Let $Y(t)$ be a subordinator. Since $Y(t)$ is nondecreasing in t, it can have at most countably many jump discontinuities in any finite time interval $[0, t]$, and at most finitely many of size $\geq \varepsilon$. (This is because these jumps cannot add up to ∞.) Also, the locations and sizes of the jumps in disjoint times intervals are independent, by the independent increments property. In fact, the jumps and increments are jointly independent across disjoint time intervals, and so if

$$X_J(t) := \sum \text{jumps up to time } t$$

then the vector-valued process $(X(t), X_J(t))$ will have stationary, independent increments, and so itself will be a Lévy process. Consequently, the process $Y(t) := X(t) - X_J(t)$ is a subordinator with no jumps. By Corollary 4.1, there is a constant $C \geq 0$ such that $Y(t) = Ct$.

To complete the proof, it suffices to show that the process $X_J(t)$ has the form (5.6) for a Poisson point process N that satisfies the hypotheses of Proposition 5.1. Since $X_J(t)$ is defined to be the sum of the jumps up to time t, it is automatically of the form (5.6); the only thing to be shown is that the point process N of jumps is a *Poisson* point process with an intensity measure that satisfies (5.8). This is (at least for now) left as an *exercise.* □

6. Appendix: Tools from Measure Theory

6.1. **Measures and continuous functions.** Many distributional identities can be interpreted as assertions that two finite measures are equal. For instance, in the identity (3.5), for each fixed set A the right and left sides both (implicitly) define measures on the Borel subsets $B$of $\mathbb{R}^k$. The next proposition gives a useful criterion for determining when two finite Borel measures are equal.

Proposition 6.1. *Let μ and ν be finite Borel measures on $\mathbb{R}^k$ such that for every bounded, continuous function $f : \mathbb{R}^k \to \mathbb{R}$,*

$$\int f\,d\mu = \int f\,d\nu. \tag{6.1}$$

*Then $\mu = \nu$, and (**??**) holds for every bounded* Borel measurable *function $f : \mathbb{R}^k \to \mathbb{R}$.*

Proof. It suffices to prove that the identity (**??**) holds for every function of the form $f = \mathbf{1}_B$ where $B \in \mathscr{B}_k$, because if it holds for such functions it will then hold for simple functions f, and consequently for all nonnegative functions by monotone convergence. Thus, we must show that

$$\mu(B) = \nu(B) \quad \forall\, B \in \mathscr{B}_k. \tag{6.2}$$

The collection of Borel sets B for which (**??**) holds is a $\lambda-$system, because both μ and ν are countably additive. Hence, it suffices to prove that (**??**) holds for all sets B in a $\pi-$system that generates the Borel $\sigma-$algebra. One natural choice is the collection $\mathscr{R}$ of *open* rectangles with sides parallel to the coordinate axes, that is, sets of the form

$$R = J_1 \times J_2 \times \cdots \times J_k$$

where each J_i is an open interval in $\mathbb{R}$. To show that (**??**) holds for $B = R$, I will show that there are bounded, continuous functions f_n such that $f_n \uparrow \mathbf{1}_R$ pointwise; since (**??**) holds for each $f = f_n$, the dominated convergence theorem will then imply that (**??**) holds for $f = \mathbf{1}_R$. To approximate the indicator of a rectangle R from below by continuous functions f_n, set

$$f_n(x) = \min(1, n \times \text{distance}(x, R^c)).$$

□

6.2. **Poisson Random Measures.** To prove that a given collection of random variables $\{N(B)\}_{B\in\mathscr{B}}$ is a Poisson point process, one must show that for each B such that $\mu(B) < \infty$ the random variable $N(B)$ has a Poisson distribution. Often it is possible to check this for sets B of a simple form, e.g., rectangles. The following proposition shows that this is enough:

Proposition 6.2. *Let $\{N(B)\}_{B\in\mathscr{B}}$ be a random measure on $(\mathbb{R}^k, \mathscr{B}_k)$ with intensity measure μ. Let $\mathscr{R} = \mathscr{R}_k$ be the collection of open rectangles with sides parallel to the coordinate axes. If*

(A) *$N(B) \sim$Poisson-$(\mu(B))$ for each $B \in \mathscr{R}$ such that $\mu(B) < \infty$, and*

(B) *$N(B_i)$ are mutually independent for all pairwise disjoint sequences $B_i \in \mathscr{R}$,*

then $\{N(B)\}_{B\in\mathscr{B}}$ is a Poisson point process.

Proof. First, it is enough to consider only the case where $\mu(\mathbb{R}^k) < \infty$. (Exercise: Check this. You should recall that the intensity measure μ was assumed to be $\sigma-$finite.) So assume this, and let $\mathscr{G}$ be the collection of all Borel sets B for which $N(B)$ is Poisson with mean $\mu(B)$. This collection is a *monotone class*[2], because pointwise limits of Poisson random variables are Poisson. Furthermore, the collection $\mathscr{G}$ contains the algebra of all

[2]A collection $\mathscr{M}$ of sets is a monotone class if it is closed under countable increasing unions and countable decreasing intersections, that is,

(1) If $A_1 \subset A_2 \subset \cdots$ are elements of $\mathscr{M}$ then so is $\cup_n A_n$; and
(2) If $B_1 \supset B_2 \supset \cdots$ are elements of $\mathscr{M}$ then so is $\cap_n B_n$.

finite unions of rectangles $R_i \in \mathscr{R}$, by hypotheses (A)-(B). Hence, by the Monotone Class Theorem, $\mathscr{G} = \mathscr{B}_k$. This proves that $N(B)$ is Poisson-$\mu(B)$ for every $B \in \mathscr{B}$.

To complete the proof we must show that for every sequence B_i of pairwise disjoint Borel sets the random variables $N(B_i)$ are mutually independent. For this it suffices to show that for any $m \geq 2$ the random variables $\{N(B_i)\}_{i \leq m}$ are independent. I will do the case $m = 2$ and leave the general case $m \geq 3$ as an exercise.

Denote by $\mathscr{A}$ the collection of all finite unions of rectangles $R \in \mathscr{R}$. If $A, B \in \mathscr{A}$ are disjoint, then $N(A)$ and $N(B)$ are clearly independent, by Hypothesis (B). Let $A, B \in \mathscr{B}$ be any two disjoint Borel sets: I will show that there exist sequences $A_n, B_n \in \mathscr{A}$ such that

(i) $\lim_{n\to\infty} \mu(A_n \Delta A) = 0$;
(ii) $\lim_{n\to\infty} \mu(B_n \Delta B) = 0$; and
(iii) $A_n \cap B_n = \emptyset$ for every $n \geq 1$.

Since $N(A_n)$ and $N(B_n)$ are independent, for each n, this will imply that $N(A)$ and $N(B)$ are independent. Recall that any Borel set can be arbitrarily well-approximated by finite unions of rectangles, so it is easy to find sequences A_n and B_n such that (i)-(ii) hold. The tricky bit is to get (iii). But note that if (i) and (ii) hold, then

$$\mu(A_n \cap B_n) \leq \mu(A_n \Delta A) + \mu(B_n \Delta B) \longrightarrow 0$$

Thus, if we replace each B_n by $B_n \setminus A_n$ then (i)- (iii) will all hold. □

Exercise 6.1. Fill in the missing steps of the proof above:

(A) Show that it is enough to consider the case where $\mu(\mathbb{R}^k) < \infty$.
(B) Show that if $X_n \perp Y_n \; \forall\, n$ and $X_n \to X$ a.s. and $Y_n \to Y$ a.s., then $X \perp Y$.
(C) Show how to extend the last part of the argument to $m \geq 3$.

UNIVERSITY OF CHICAGO, DEPARTMENT OF STATISTICS, USA