Infinitely divisible processes

Adding up the points

The third axiom for a Poisson process is that the function $A \mapsto N_A$ is (almost surely) a measure on S. Suppose that $S = \mathbb{R}$ (or some other banach space). Then we can "add up the points", that is, define the random variable Z by $Z(\omega) = \int x\, dN.$ where $N. = N.(\omega)$ is the random measure $A \mapsto N_A$. If μ is finite, then N is almost surely a finite counting measure (the Poisson process has almost surely finitely many points) and the integral is defined almost surely.

Suppose that μ is infinite. To see whether the integral is well defined, we look at the positive and negative parts and see whether they are finite. Thus we restrict for now to the case $S = \mathbb{R}^+$. Intuitively, since the measure μ is σ-finite, there are only countably many points and we may write the integral as a countable sum. Formally, write μ as the increasing limit of finite measures, to see that we really do add an increasing limit of finite sets of points. If $\mu(x, \infty) = \infty$ for any $x > 0$, then there are infinitely many points of magnitude at least x, so Z is almost surely infinite. If not, but $\mu(0, x) = \infty$, then there are infinitely many points to sum and Z may or may not be finite. The expectation of Z is $\int x\, d\mu$:

$$\mathbb{E} \int x\, dN(x) = \lim_{\epsilon \to 0} \sum_n n\epsilon\mu(n\epsilon, (n+1)\epsilon] \to \int x\, d\mu < \infty\, ;$$

if this is finite then Z is almost surely finite. If $\int_0^\epsilon x\, d\mu = \infty$, then (exercise) by souped-up Borel-Cantelli, Z is almost surely infinite.

We see therefore that the correct test is the finiteness of $\int_0^\epsilon x\, d\mu + \mu(\epsilon, \infty)$ (it does not matter what value of ϵ is chosen). There are a number of equivalent formulations; one convenient one is the finiteness of $\int x/(1+x)\, d\mu$. When this condition is satisfied, we use the phrase "adding up the points of the Poisson process" to denote the random variable Z.

The compound Possion distribution

In your homework, you have already computed the characteristic function for a compound Poisson, that is, the sum of Poisson-many IID random variables. Let us generalize this to the characteristic function for the sum of the points of a Poisson process. We build up to this as follows.

Why is the characteristic function for the Poisson what it is? Heuristic: $\mathcal{P}(\lambda)$ may be thought of as the convolution of $1/dx$-many variables that are IID $\sim \mathcal{P}(\lambda\, dx)$. Each of these characteristic functions is $1 + \lambda\, dx\, (e^{it} - 1)$, so, taking logs, the characteristic function of the $\mathcal{P}(\lambda)$ is given by a sum of $1/dx$-many terms of order dx, that is, an integral:

$$\phi(t) = \exp[\int_0^1 \lambda dx(e^{it} - 1)] = \exp[\lambda(e^{it} - 1)]\,.$$

What if we have a rate λ Poisson process but each point that occurs counts as 2? If $Z \sim \mathcal{P}(\lambda)$ then this new variable Y is distributed as $2Z$, so we can see immediately that $\phi_Y(t) = \phi_Z(2t)$. But another way to see this is to note that the infinitesimal contribution becomes $\lambda\, dx\, (e^{2it} - 1)$, leading to $\log \phi_Y(t) = \lambda(e^{2it} - 1)$.

This is better because it generalizes. Suppose each point has a value attached and these values are IID $\sim \mu$. Then in interval dx we have a $1 - \lambda\, dx$ probability of adding zero and a $\lambda\, dx$ probability of adding a fresh pick from μ, so $\mathbb{E}e^{itX} = 1 + \lambda dx(\mathbb{E}_\mu e^{itY} - 1)$, and taking the integral of the log gives

$$\phi(t) = \exp[\lambda(\phi_\mu(t) - 1)]\,. \tag{1}$$

Thus we recover the result of a previous homework problem.

Our final generalization is to sum the points of a Poisson process with intensity μ on $\mathbb{R}$. We assume that the integral is almost surely well defined, that is, $\int |x|/(1 + |x|)\, d\mu < \infty$. The interval $[x, x + dx]$ now contributes a summand of $\mu(x, x + dx)e^{itx} - 1$ to $\log \phi(t)$, yielding

$$\log \phi(t) = \int \left(e^{itx} - 1\right)\, d\mu(x)\,. \tag{2}$$

This was not rigorous, so let us now prove it is correct.

Theorem 1 *Let μ be a measure on $\mathbb{R}$ with $\int \frac{|x|}{1+|x|}\, d\mu < \infty$. Then the sum of the points of a Poisson process with intensity μ on $\mathbb{R}$ is well defined, and has characteristic function given by (2) which is also well defined.*

Proof: Since $e^{itx} - 1 = O(|x|/(1 + |x|)$ on all of $\mathbb{R}$, we see that the integral is well defined. If μ is finite with mass λ, the second construction of the Poisson process constructs the sum of the points as the sum of $\mathcal{P}(\lambda)$ many IID variables with distribution $\mu/||\mu||$. Plugging into (1) yields

$$\begin{aligned} \log \phi(t) &= ||\mu||(\phi_{\mu/||\mu||}(t) - 1) \\ &= ||\mu|| \int (e^{itx} - 1)\, d(\mu/||\mu||)(x) \end{aligned}$$

which agrees with (2). If μ is infinite, then Z is the limit in distribution of the sum, Z_ϵ, of the points in the complement of $[-\epsilon, \epsilon]$; the measure $\mu_{[-\epsilon,\epsilon]^c}$ is finite, so its characteristic function is

$$\log \phi_\epsilon(t) = \int_{[-\epsilon,\epsilon]^c} \left(e^{itx} - 1\right) \, d\mu(x) \, .$$

The integral is dominated by $\frac{C|x|}{1+|x|} \, d\mu(x)$ which is finite by assumption, so by dominated convergence, the right-hand side converges to $\int (e^{itx} - 1) \, d\mu(x)$. The assumptions also imply that the total positive and total negative increments are almost surely finite, hence $Z_\epsilon \to Z$ almost surely as $\epsilon \to 0$, hence $\phi_\epsilon(t) \to \phi(t)$, giving the desired conclusion. □

Infinitely divisible distributions

An infinitely divisible distribution is defined in Durrett to be a limit of sums of n IID random variables as $n \to \infty$. It is easy to see this is equivalent to the stronger condition that it be equal to a sum of n IID's for any n. Proof: if it is the limit of sums of $2n$ IID's, then grouping the first and the second halves, it is the sum of two IID's. Same works for any finite number (see proof of II.8.1 in Durrett).

Lévy processes

Why do we care about this? One reason is that from a process point of view, these are what we need to define TIME-HOMOGENEOUS LEVY PROCESSES: are processes in which $X_{s_1}, X_{s_2} - X_{s_1}, X_{s_3} - X_{s_2}, \ldots$ are IID for any s. Examples: (1) Brownian motion; (2) the Poisson process. These two processes are very different, e.g., B_t is a.s. continuous, while N_t is a.s. a counting step function. We'd like to know more of these; for example, in modeling it will make sense to interpolate between these two extremes. If F is infinitely divisible, then can define a THLP on $[0, 1]$ as follows. For each k, define $\{X_{j/2^k} : 1 \leq j \leq 2^k\}$ by letting $\{X_{(j+1)/2^k} - X_{j/2^k}\}$ be IID's whose sum has distribution F. The joint distributions are consistent for different k, so there is a limiting distribution $\{X_t : \exists j, k\ t = j/2^k\}$. It is not clear that this process will extend to a nice process $\{X_t : t \in [0,1]\}$ but any THLP on $[0,1]$ will specialize to such a process on the dyadic rationals.

Let us see how to use summing points of a Poisson process to generalize the previous two examples, almost universally. Let μ be a measure satisfying the previous integrability condition $\int |x|/(1+|x|)\, d\mu < \infty$. Consider a Poisson process on $\mathbb{R} \times \mathbb{R}^+$ with intensity $\mu \times dy$. Fix $\lambda > 0$ and for measurable $A \subseteq \mathbb{R}$ define $N_A = N_{A\times[0,\lambda]}$. That is, we throw away all the points in the half plane above $y = \lambda$ and then project down onto the x-axis. Clearly, this constructs a Poisson process of intensity $\lambda\mu$. Furthermore, it does this simultaneously for all $\lambda > 0$. Let Z_λ denote the sum of the points. Then $\{Z_\lambda : \lambda > 0\}$ is a THLP, and the distribution of any of these, say Z_1, is infinitely divisible.

Compensated increments

Having introduced the idea of a 2-dimensional Poisson process in order to couple Poisson processes with intensity $\lambda\mu$ for all λ, let us now go back to summing the ordinates of Poisson point processes on $\mathbb{R}$.

Idea: if $\{X_j\}$ are IID with $\mathbb{P}(X_j = 1/N) = \mathbb{P}(X_j = 0) = 1/2$, we can only add up order N of these if we want the sum to remain tight, but if $\mathbb{P}(X_j = 1/N) = \mathbb{P}(X_j = -1/N) = 1/2$, we can add up order N^2 of them.

This suggests that we might be able add more increments (weaken the condition $\int_\epsilon^\epsilon |x|\, d\mu < \infty$) if we sum symmetric intensities over symmetric intervals. We can and will do that, but it is more general simply to sum compensated random variables.

Compensated Poisson

What if intstead of $\mathbb{P}(dX = 1) = 1 - \mathbb{P}(dx = 0) = \lambda\, dt$, we make the increment mean zero:

$$P(dX = 1 - \lambda\, dt) = 1 - P(X = -\lambda\, dt) = \lambda dt\,.$$

Summing the increments gives process $N_t - \lambda\, t$ (though we prefer to think of the $-\lambda t$ as built of increments of $-\lambda dt$. Recentering, of course, causes a multiplication by $\exp(-i\lambda t)$ but again, another way to write this is in the integral,

$$\phi(t) = \exp[\lambda(e^{it} - 1 - it)]\,.$$

Similarly, one can compensate a compound Poisson and get

$$\phi(t) = \exp[\int (e^{itX} - 1 - itX) d\mu(X)]. \tag{3}$$

Note that when μ is a finite measure and X is integrable, this is still the same as multiplying the noncompensated Poisson characteristic function by $e^{-it\mathbb{E}X}$.

Infinitely many points

Now how many increments can we sum and still remain tight? We can view this both analytically and probabilistically. The probabilistic view is that if the contribution of the points in an interval $[x, x + dx]$ has variance $V\,dx$ AND MEAN ZERO, then we should obtain a tight limit (in fact a limit in L^2) if V is integrable near zero. The integrability condition at infinity is still $\mu(a, \infty) < \infty$, guaranteeing to only finitely many jumps beyond any given size. The corresponding analytic fact is that $|e^{it} - 1 - it| = \Theta(t^2)$ as $t \to 0$, hence the integral in (3) exists near zero if and only if $\int_{-\epsilon}^{\epsilon} x^2\, d\mu < \infty$. In that case, the integral is the limit of integral over $(-\epsilon, \epsilon)^c$, so by the continuity theorem, summing all the compensated points in $[-a, a]$ is the (well defined) distributional limit as $\delta, \epsilon \to 0$ of summing the compensated points in $[-a, a] \setminus (-\epsilon, \delta)$. When μ is symmetric, if we like, we can take $\epsilon = \delta$ and forget the compensation.

Putting together this integrability condition near zero with $\mu(-\epsilon, \epsilon)^c < \infty$ gives a single equivalent integrability condition, namely $\int x^2/(1 + x^2)\, d\mu(x) < \infty$. It turns out that summing in this way this gives all infinitely divisible distributions!

Theorem 2 (Lévy-Khinchin) *Let μ be any measure with $\mu(\{0\}) = 0$ and $\int x^2/(1 + x^2)\, d\mu(x) < \infty$. Then for any $c \in \mathbb{R}$ and $\sigma^2 \geq 0$, there is an infinitely divisible measure whose characteristic function ϕ satisfies*

$$\log \phi(t) = ict - \frac{1}{2}\sigma^2 t^2 + \int \left[e^{itx} - 1 - \frac{itx}{1 + x^2} \right] d\mu(x)\,.$$

Furthermore, this gives all infinitely divisible distributions.

(SKETCH OF HALF OF) PROOF: First let's explain the formula. The first two terms say we can add a normal with any mean and variance (since these are obviously infinitely divisible; for another interpretation of this, stay tuned). If we compensate in $[x, x + dx]$ by $-x/(1 + x^2)\, d\mu(x)$ then the mean is no longer zero but is $x(1 - 1/(1 + x^2))\, d\mu(x)$. But $x - x/(1 + x^2) = O(x^3)$ near zero, so the variance condition, i.e., that $x^2\, d\mu$ be integrable near zero, more than implies that the means will be summable. Since $x/(1+x^2)$ is bounded, we also have that the total compensation in $[-a, a]^c$ will be finite (since $\mu([-a, a]^c < \infty)$; since the number of points there is finite too, the sum of the compensated contributions is almost surely finite. Thus the strangely compensated variables have a well defined sum, and we have already seen that the characteristic function is the one given in the theorem. For the other direction (that this is all), see the references given in Durrett. □

Finite variance case

The small contributions (the compensated jumps in $[-a, a]$) have to have finite total variance in order for the sum of the compensated jumps to be defined. If $\int_{[-a,a]^c} x^2\, d\mu$ happens to be finite, then the variance of the large contributions is finite as well. In this case, it is valid and simpler to compensate correctly, by $x d\mu$ rather than $x/(1 + x^2)\, d\mu$. We may now write

$$\log \phi(t) = \int \left(e^{itx} - 1 - itx\right)\, d\mu(x)$$

for μ such that $\int x^2\, d\mu < \infty$. Writing ν for the measure $x^2\, d\mu$, this becomes

$$\log \phi(t) = \int \left(\frac{e^{itx} - 1 - itx}{x^2}\right)\, d\nu(x)$$

where ν runs over arbitrary measures of finite total mass. [This is Durrett's (8.5).] If we define the integrand to be $-t^2/2$ at $x = 0$, then a point mass σ^2 in ν at zero yields a summand $-\sigma^2 t^2/2$ in $\log \phi$, thus we see this formula is broad enough to include a centered normal term; adding ict for $c \in \mathbb{R}$ gives a formula for the general finite-variance infinitely divisible distribution.