Statistics and Probability Letters 232 (2026) 110653

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The law of thin processes: A law of large numbers for point processes

Matthew Aldridge

School of Mathematics, University of Leeds, Leeds, LS2 9JT, UK

ARTICLE INFO

Keywords:
Point processes
Poisson process
Law of large numbers

ABSTRACT

If you take a superposition of n IID copies of a point process and thin that by a factor of $1/n$, then the resulting process tends to a Poisson process as $n \to \infty$. We give a simple proof of this result that highlights its similarity to the law of large numbers and to the law of thin numbers of Harremoës et al.

1. Theorems

The law of large numbers is the following result.

Theorem 1 (*Law of Large Numbers*). *Let X be a real-valued random variable with finite expectation $\mu = \mathbb{E}X$. Let $X_1, X_2, \ldots$ be IID copies of X. Then the scaled sum*

$$\frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

tends in distribution to a point mass at μ as $n \to \infty$.

(The weak law of large numbers is more often presented as concerning convergence in probability to μ, but convergence in probability to a constant and convergence in distribution to a point mass at that constant are equivalent.)

The steps here are: first, we sum n IID copies of a random variable; second, we scale that sum by a factor of $1/n$; then we converge to a distribution that depends only on the first moment μ of the random variable we started with.

Harremoës, Johnson, and Kontoyiannis (2010) give a similar result for discrete random variables that take values in the non-negative integers, where the limit theorem remains a statement about the non-negative integers. While the law of large numbers would still apply for such a random variable, the act of scaling the sum by $1/n$ takes us outside the non-negative integers, and the limiting distribution of a point mass at μ is also not supported on the non-negative integers (unless μ itself happens to be an integer). Harremoës et al. make two changes to the law of large numbers to remedy this. The first change is that the scaling operation is replaced with *thinning*. To get the thinning $p \circ X$ of a non-negative integer random variable X, one thinks of X as representing a number of items, each of which is independently kept with probability p and removed with probability $1 - p$; more formally, the conditional distribution of $p \circ X$ given X is the binomial distribution $\mathrm{Bin}(X, p)$. The second change is that the limiting distribution of the point mass at μ is replaced by a Poisson distribution with rate μ. Harremoës et al. call this result 'the law of thin numbers'.

Theorem 2 (*Law of Thin Numbers; Harremoës et al., 2010*). *Let X be a random variable on the non-negative integers with finite expectation $\mu = \mathbb{E}X$. Let $X_1, X_2, \ldots$ be IID copies of X. Then the thinned sum*

$$\frac{1}{n} \circ (X_1 + X_2 + \cdots + X_n)$$

tends in distribution to a Poisson distribution with rate μ as $n \to \infty$.

E-mail address: m.aldridge@leeds.ac.uk.

https://doi.org/10.1016/j.spl.2026.110653
Received 21 March 2025; Received in revised form 19 August 2025; Accepted 13 January 2026
Available online 14 January 2026

Note the same essential steps as the law of large numbers: first, we sum n IID copies of a discrete random variable; second, we thin that sum by a factor of $1/n$; then we converge to a distribution that depends only on the first moment μ of the random variable we started with.

In this note, we show that there exists a law of large numbers for point processes too. Consider a point process ξ. First, take a superposition of n IID copies of ξ. We think of a point process as a random counting measure, so the superposition is a sum of n IID copies of the random measure. Second, thin that superposition by a factor of $1/n$. To get the thinning $p\circ\xi$ of a point process ξ, we independently keep each point of ξ with probability p and delete it with probability $1-p$. Then, as $n\to\infty$, we get a Poisson process with the same intensity measures the original point process. Following Harremoës et al. we propose the name 'the law of thin processes'.

We largely use notation and terminology from the book of Daley and Vere-Jones (2008).

Theorem 3 (*Law of Thin Processes*)**.** *Let ξ be a point process on a complete separable metric space $\mathcal{X}$ with boundedly finite intensity measure μ, where $\mu(A)=\mathbb{E}\xi(A)$. Let $\xi_1,\xi_2,\dots$ be IID copies of ξ. Then the thinned superposition*

$$\frac{1}{n}\circ(\xi_1+\xi_2+\cdots+\xi_n)$$

tends weakly to a Poisson process on $\mathcal{X}$ with intensity measure μ as $n\to\infty$.

Again, the result has the same steps as the laws of large and thin numbers: first, we superimpose n IID copies of a point process; second, we thin that superposition by a factor of $1/n$; then we converge to a law that depends only on the first moment measure μ of the point process we started with.

Although I have not managed to find the law of thin processes in the literature exactly as written above, there are many results on the same lines. Serfozo (1984) has a limit theorem for more general concepts of thinning a point process, and the law of thin processes is essentially a special case of that result. Exercise 5.3 of Last and Penrose (2018) invites the reader to find ξ such that a thinned superposition of IID copies $\frac{1}{n}\circ(\xi_1+\cdots+\xi_n)$ has the same distribution as ξ itself: this is any Poisson process. There are two results in Daley and Vere-Jones (2008) for point processes in $\mathbb{R}$ that replace one of the two operations (superposition or thinning) with dilation of the real line: Proposition 11.2.VI keeps the superposition of n IID copies of a real-valued point process, but replaces the thinning with a dilation of space by a factor of n; while Proposition 11.3.I replaces the superposition with a contraction of space by a factor of $1/n$ but keeps the thinning. Both results show convergence to a Poisson process, under a stationarity or 'weak stationarity' assumption. Moreover, the law of thin processes can presumably be shown as a corollary of more general (and complicated) Poisson limit theorems, such as those of Daley and Vere-Jones (2008, Proposition 11.2.V) or Kallenberg (2017, Corollary 4.41).

However, our emphasis in this note is not on novelty of the result itself, nor on presenting Poisson convergence in greatest generality, but rather on giving a simple proof of Theorem 3 that exactly reflects standard proofs for the laws of large and thin numbers.

2. Proofs

2.1. Laws of large and thin numbers

We start by sketching the proofs of the laws of large and thin numbers, because their structures – which are extremely similar – will show us the route to take to prove the law of thin processes and will highlight the similarities between the three results.

The proof of the law of large numbers will use the Laplace transform $L_X(u)=\mathbb{E}\mathrm{e}^{-uX}$ for u a real number in some neighbourhood U of 0. (One could just as well use the moment generating function $M_X(u)=L_X(-u)=\mathbb{E}\mathrm{e}^{uX}$, but we keep the minus sign here for consistency with later proofs. For the purpose of this proof we will assume that $L_X(u)$ does exist within such a U, but if convergence is an issue, one can instead use the characteristic function $\Phi_X(u)=L_X(-\mathrm{i}u)=\mathbb{E}\mathrm{e}^{\mathrm{i}uX}$, which exists for all $u\in\mathbb{R}$.)

The Laplace transform has the following three crucial properties:

1. Behaviour under independent sums: If X and Y are independent, then $L_{X+Y}(u)=L_X(u)\,L_Y(u)$ for all $u\in U$.
2. Behaviour under scaling: $L_{aX}(u)=L_X(au)$ for all $a\in\mathbb{R}$ and $u\in U$.
3. Behaviour with limits: If $L_{X_n}(u)\to L_X(u)$ for all $u\in U$ as $n\to\infty$, then $X_n\to X$ in distribution.

Proof of Theorem 1. The Laplace transform of X is

$$L_X(u)=\mathbb{E}\mathrm{e}^{-uX}=\mathbb{E}\sum_{j=0}^{\infty}\frac{(-uX)^j}{j!}=\sum_{j=0}^{\infty}\frac{(-1)^j}{j!}(\mathbb{E}X^j)\,u^j=1-\mu u+o(u).$$

Writing

$$Y_n=\frac{1}{n}(X_1+X_2+\cdots+X_n)$$

for the scaled sum, we have, using facts 1 and 2 above,

$$L_{Y_n}(u)=L_X\left(\frac{1}{n}u\right)^n=\left(1-\frac{\mu u}{n}+o\left(\frac{1}{n}\right)\right)^n\to\mathrm{e}^{-\mu u}$$

as $n\to\infty$. But this is the Laplace transform of a point mass at μ, so by fact 3, Y_n converges in distribution to that point mass as $n\to\infty$. □

The proof of the law of thin numbers will use the alternate probability generating function $A_X(u) = \mathbb{E}(1-u)^X$ for $u \in [0,2]$. The alternate probability generating function has the following three crucial properties, equivalent to the three properties of the Laplace transform:

1. Behaviour under independent sums: If X and Y are independent, then $A_{X+Y}(u) = A_X(u)\,A_Y(u)$ for all $u \in [0,2]$.
2. Behaviour under thinning: $A_{p\circ X}(u) = A_X(pu)$ for all $p \in [0,1]$ and $u \in [0,2]$.
3. Behaviour with limits: If $A_{X_n}(u) \to A_X(u)$ for all $u \in [0,2]$ as $n \to \infty$, then $X_n \to X$ in distribution.

Fact 2 is why we prefer the alternate probability generating function to the more common probability generating function $G_X(u) = \mathbb{E}u^X = A_X(1-u)$, since that has the more awkward expression $G_{p\circ X}(u) = G_X(1-p+pu)$.

The following proof of the law of thin numbers is essentially that given in Jørgensen and Kokonendji (2016, Proposition 2.1).

Proof of Theorem 2. The alternate probability generating function of X is

$$A_X(u) = \mathbb{E}(1-u)^X = \mathbb{E}\sum_{j=0}^{\infty}\binom{X}{j}(-u)^j = \sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\,\mathbb{E}(X)_j\,u^j = 1 - \mu u + o(u),$$

where $\mathbb{E}(X)_j = \mathbb{E}X(X-1)\cdots(X-j+1)$ are the factorial moments. Writing

$$Y_n = \frac{1}{n}\circ(X_1 + X_2 + \cdots + X_n)$$

for the thinned sum, we have, using facts 1 and 2 above,

$$A_{Y_n}(u) = A_X\left(\frac{1}{n}\,u\right)^n = \left(1 - \frac{\mu u}{n} + o\left(\frac{1}{n}\right)\right)^n \to \mathrm{e}^{-\mu u}$$

as $n \to \infty$. But this is the alternate probability generating function of a Poisson distribution with rate μ, so by fact 3, Y_n converges in distribution to that Poisson distribution as $n \to \infty$. $\square$

2.2. *Law of thin processes*

To prove the law of thin processes (Theorem 3), we use what we shall call the *alternate probability generating functional*

$$A_\xi(u) = \mathbb{E}\exp\left(\int_{\mathcal{X}}\log\big(1-u(x)\big)\,\xi(\mathrm{d}x)\right)$$

for $u \in \mathcal{U}$, where $\mathcal{U} = \mathcal{U}(\mathcal{X})$ is the set of functions $u : \mathcal{X} \to [0,1]$ that are zero outside a bounded set. Because the point process is almost surely finite on the set where u does not vanish, we can write

$$A_\xi(u) = \mathbb{E}\prod_{X\in\xi}\big(1-u(X)\big),$$

where the product is taken over the points X of the point process ξ.

The alternate probability generating functional again has three crucial properties, which mirror those we saw for the Laplace transform and the alternate probability generating function in the two earlier proofs:

1. Behaviour under independent superpositions: If ξ and η are independent, then $A_{\xi+\eta}(u) = A_\xi(u)\,A_\eta(u)$ for all $u \in \mathcal{U}$ (Daley and Vere-Jones, 2008, Proposition 9.4.IX).
2. Behaviour under thinning: $A_{p\circ\xi}(u) = A_\xi(pu)$ for all $p \in [0,1]$ and $u \in \mathcal{U}$ (Daley and Vere-Jones, 2008, equation (11.3.2)).
3. Behaviour with limits: If $A_{\xi_n}(u) \to A_\xi(u)$ for all $u \in \mathcal{U}$ as $n \to \infty$, then $\xi_n \to \xi$ weakly (Daley and Vere-Jones, 2008, Proposition 11.1.VIII).

Again, fact 2 is why we prefer the alternate probability generating functional to the more common probability generating functional

$$G_\xi(u) = \mathbb{E}\exp\left(\int_{\mathcal{X}}\log u(x)\,\xi(\mathrm{d}x)\right) = \mathbb{E}\prod_{X\in\xi}u(X) = A_\xi(1-u),$$

which has the awkward expression $G_{p\circ\xi}(u) = G_\xi(1-p+pu)$, or the Laplace functional

$$L_\xi(u) = \mathbb{E}\exp\left(-\int_{\mathcal{X}}u(x)\,\xi(\mathrm{d}x)\right) = \mathbb{E}\prod_{X\in\xi}\mathrm{e}^{-u(X)} = A_\xi(1-\mathrm{e}^{-u}),$$

for which $L_{p\circ\xi}(u) = L_\xi(-\log(1-p+p\mathrm{e}^{-u}))$.

The final preparatory step we need is a result that writes the alternate probability generating functional in terms of the factorial moment measures $m_{(j)}$, in the same way as we earlier wrote the Laplace transform in term of the moments $\mathbb{E}X^j$ and the alternate probability generating function in terms of the factorial moments $\mathbb{E}(X)_j$. Informally, we would expect

$$A_\xi(u) = 1 + \sum_{j=1}^{\infty}\frac{(-1)^j}{j!}\int_{\mathcal{X}^j}u(x_1)\cdots u(x_j)\,m_{(j)}(\mathrm{d}x_1\times\cdots\times\mathrm{d}x_j).$$

However, because we now want to give a formal proof, we have to be careful about convergence issues that mean the summands on the right-hand side may not exist or the sum may not converge (see Last and Penrose (2018, Lemma 4.11)). What we do have if the following: if ξ is such that the kth factorial moment measure $m_{(k)}$ exists, then

$$A_\xi(pu) = 1 + \sum_{j=1}^{k} \frac{(-p)^j}{j!} \int_{\mathcal{X}^j} u(x_1)\cdots u(x_j)\, m_{(j)}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_j) + o(p^k)$$

for $u \in \mathcal{U}$ as $p \to 0$ (Daley and Vere-Jones, 2008, Proposition 9.5.VI). The precise definition of the factorial moment measures is not important here, as we will only need the $k = 1$ case,

$$A_\xi(pu) = 1 - p \int_{\mathcal{X}} u(x)\, \mu(\mathrm{d}x) + o(p), \tag{1}$$

where the first factorial moment measure $m_{(1)} = \mu$ is the simply the intensity measure $\mu(A) = \mathbb{E}\xi(A)$ of ξ.

We are now ready to prove the law of thin processes by following exactly the steps of the two earlier proofs.

Proof of Theorem 3. Writing

$$\eta_n = \frac{1}{n} \circ (\xi_1 + \xi_2 + \cdots + \xi_n)$$

for the thinned superposition, we have

$$A_{\eta_n}(u) = A_\xi\left(\frac{1}{n}u\right)^n = \left(1 - \frac{1}{n}\int u(x)\,\mu(\mathrm{d}x) + o\left(\frac{1}{n}\right)\right)^n$$

as $n \to \infty$, where the first equality is from facts 1 and 2 above, and the second equality is the result (1). Then

$$A_{\eta_n}(u) = \left(1 - \frac{1}{n}\int u(x)\,\mu(\mathrm{d}x) + o\left(\frac{1}{n}\right)\right)^n \to \exp\left(-\int u(x)\,\mu(\mathrm{d}x)\right).$$

But this is the alternate probability generating functional of a Poisson process with intensity measure μ (Daley and Vere-Jones, 2008, Example 9.4(c)), so by fact 3, η_n converges weakly to that Poisson process as $n \to \infty$. □

Data availability

No data was used for the research described in the article.

References

Daley, D.J., Vere-Jones, D., 2008. An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure, second ed. In: Probability and its Applications, Springer, New York, http://dx.doi.org/10.1007/978-0-387-49835-5.

Harremoës, P., Johnson, O., Kontoyiannis, I., 2010. Thinning, entropy, and the law of thin numbers. IEEE Trans. Inform. Theory 56 (9), 4228–4244. http://dx.doi.org/10.1109/TIT.2010.2053893.

Jørgensen, B., Kokonendji, C.C., 2016. Discrete dispersion models and their Tweedie asymptotics. 100, (1), pp. 43–78. http://dx.doi.org/10.1007/s10182-015-0250-z,

Kallenberg, O., 2017. Random Measures, Theory and Applications. In: Probability Theory and Stochastic Modelling, vol. 77, Springer, Cham, http://dx.doi.org/10.1007/978-3-319-41598-7.

Last, G., Penrose, M., 2018. Lectures on the Poisson Process. In: Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, Cambridge.

Serfozo, R., 1984. Thinning of cluster processes: convergence of sums of thinned point processes. Math. Oper. Res. 9 (4), 522–533. http://dx.doi.org/10.1287/moor.9.4.522.