The Annals of Probability
1983, Vol. 11, No. 4, 904–908

RAMSEY'S THEOREM AND POISSON RANDOM MEASURES

BY TIMOTHY C. BROWN AND JOSEPH KUPKA

Monash University

Prékopà's theorem gives a qualitative sufficient condition for a completely random point process to be Poisson. A generalization of this theorem is presented. The proof is elementary and uses a combinatorial principle known as Ramsey's theorem.

1. Introduction. It is well known that the Poisson and exponential distributions may be derived from purely qualitative assumptions, and they perhaps owe their ubiquity to the fact that these assumptions appear eminently reasonable in many applications. In this note we shall show that the Poisson distribution arises, in the context of integer-valued completely random measures, as a consequence of a simple combinatorial condition upon the random measure. The first theorem of this sort was due to Prékopà [9, Theorem 1, page 155]. It requires a more restrictive combinatorial condition than our own, and it also makes use of a special assumption about the underlying probability measure space. Kallenberg [4, Corollary 7.4, page 48] and Jagers [3, Theorem 4, page 215] replace this special assumption with topological conditions upon the phase space. As Ripley [10, page 983] points out, it is inconvenient to have to satisfy such topological conditions in certain problems of stochastic geometry, and they happily turn out to be unnecessary. Previous authors have also attached great importance to singleton subsets of the phase space. In the event these need play no part in the derivation of the Poisson distribution.

Our proof will be completely elementary. We shall use only the basic convergence theorems of measure theory together with a simple combinatorial principle known as Ramsey's theorem (see Section 2). The general theory spawned by this result, which is called *Ramsey theory*, has already been used extensively in the study of Banach spaces [8] and is beginning to make its influence felt in general measure theory [6]. We anticipate that it will soon find other applications in probability.

Sections 3 and 4 comprise modifications of standard arguments in random measure theory. Since these arguments are brief, we have presented them in full. This serves to make the exposition self-contained, and it also clarifies how Ramsey's theorem fits naturally into the overall development. The Poisson distribution is obtained in Section 5 by a technique which is new even in the topological setting, and Section 6 briefly explores the ties between our combinatorial condition and the "nonatomicity" of the random measure.

Before we turn to the main theorem, let us review the definition of a random measure. Let $(\Omega, \mathscr{F}, P)$ be a probability measure space, let H be a nonempty set, and let $\mathscr{D}$ be a δ-ring of subsets of H (that is, $\mathscr{D}$ is a ring which is closed under countable intersections). When $H = \mathbb{R}^k$ for some $k = 1, 2, \cdots$, the bounded Borel sets constitute a natural choice for $\mathscr{D}$. A *random measure defined on* $(\Omega, \mathscr{F}, P)$ *with phase space* $(H, \mathscr{D})$ is a function Φ which assigns to each point $\omega \in \Omega$ a nonnegative, countably additive measure $\Phi(\cdot)(\omega)$ on $\mathscr{D}$ in such a way that, for every set $D \in \mathscr{D}$, the function $\Phi(D)$ is a random variable (that is, $\Phi(D)$ is $\mathscr{F}$-measurable). Following Kingman [5, page 60], we declare Φ to be *completely random* if, for every choice of pairwise disjoint sets $D_1, D_2, \cdots, D_n \in \mathscr{D}$, the random variables $\Phi(D_1), \Phi(D_2), \cdots, \Phi(D_n)$ are mutually independent. (Useful minimal conditions for the existence of a completely random measure are provided by [7, Theroem 13.3, page 16].) We shall say that a set $D \in \mathscr{D}$ is *small* (with respect to Φ) if $\Phi(D) \leq 1$, and if $\Phi(D) = 0$ a.s. This definition is mainly intended for the case where Φ is *integer-valued*, that is, where $\Phi(D)(\omega)$ is an integer for all $\omega \in \Omega$, and for all $D \in \mathscr{D}$.

Received August 1982.
AMS 1980 *subject classifications.* 60G57, 05C55.
Key words and phrases. Poisson random measures, Prékopà's theorem, Ramsey's theorem.

We shall now state the principal result of this paper.

1.1. THEOREM. *Let Φ be an integer-valued completely random measure defined on $(\Omega, \mathscr{F}, P)$ with phase space $(H, \mathscr{D})$. Assume that*:

(1.2) *For every set $D \in \mathscr{D}$, there exists a countable family $\mathscr{B} \subseteq \mathscr{D}$ such that $D \subseteq \cup \mathscr{B}$, and such that, for all $x \in D$, the set $\cap \{B \in \mathscr{B} : x \in B\}$ is small.*

Then $\Phi(D)$ has a Poisson distribution for all $D \in \mathscr{D}$.

The traditional assumption which has been used in place of (1.2) is that singleton sets belong to $\mathscr{D}$ and are small. This assumption implies condition (1.2) in all known special cases of Theorem 1.1. For example, when H is a metric space, one may obtain (1.2) with the mild additional requirement that every set $D \in \mathscr{D}$ can be *countably* partitioned by sets in $\mathscr{D}$ of arbitrarily small diameter. In particular, such a partitioning is possible in a separable metric space, provided only, for example, that open sets are *locally measurable* (that is, $U \cap D \in \mathscr{D}$ for every set $D \in \mathscr{D}$, and for every open set U).

Although the phase space H is undeniably important in applications, it will serve mainly as a "security blanket" in the present discussion. The following condition is equivalent to (1.2), it makes sense when $(H, \mathscr{D})$ is replaced by an abstract δ-ring, and our proof of Theorem 1.1 carries over to this abstract setting.

(1.3) For every set $D \in \mathscr{D}$, there exists a countable *field* $\mathscr{B} \subseteq \mathscr{D}$ of subsets of D such that the intersection of every maximal filter in $\mathscr{B}$ is small.

When H is present, the wording in (1.2) is clearly to be preferred because of its greater simplicity.

We shall divide up the proof of Theorem 1.1 into labelled pieces. It is possible to read Section 5 prior to Sections 2, 3, and 4.

2. The fundamental sequence of partitions and Ramsey's theorem. The first step is to convert assumption (1.2) into a more workable condition. Write $\mathscr{B} = \{B_n\}_{n=1}^{\infty}$, and, for $n = 1, 2, \cdots$, let $\mathscr{P}_n$ be the partition of D which is determined by $B_1, \cdots, B_n$. (Thus, $\mathscr{P}_1 = \{D \cap B_1, D \cap B_1^c\}$, and so on.) Then the following facts are clear.

(2.1) For all n, the partition $\mathscr{P}_{n+1}$ refines the partition $\mathscr{P}_n$.

(2.2) If $\mathscr{C} \subseteq \mathscr{D}$ is countable, and if $\mathscr{C} \cap \mathscr{P}_n \neq \phi$ for infinitely many n, then $\cap \mathscr{C}$ is small.

Any sequence of finite partitions $\mathscr{P}_n \subseteq \mathscr{D}$ of a set $D \in \mathscr{D}$ which satisfies conditions (2.1) and (2.2) will be called a *fundamental sequence of partitions* for D.

Condition (2.2) delineates a "local" phenomenon. In order to derive "global" conclusions from it, we shall invoke the following combinatorial result. It is a special case of Ramsey's theorem, and its short, simple proof is to be found at the *top* of page 383 of [8].

2.3. THEOREM. *Let each pair $\{i, j\}$ of distinct positive integers be colored either "red" or "blue" in an arbitrary fashion. Then there exists a strictly increasing sequence $\{n_i\}_{i=1}^{\infty}$ of positive integers such that all of the pairs from this sequence have the same color.*

Consider now the following "local" fact. If $\{D_n\}_{n=1}^{\infty}$ is a decreasing sequence of sets in $\mathscr{D}$ whose intersection D is small, then we have $\Phi(D_n) \downarrow \Phi(D) \leq 1$ by the countable additivity of Φ. Since the $\Phi(D_n)$ are integer-valued, it follows, for each point $\omega \in \Omega$, that $\Phi(D_n)(\omega) \leq 1$ for all n sufficiently large. We shall now use Theorem 2.3 to convert this observation into the following "global" fact about a fundamental sequence $\{\mathscr{P}_n\}_{n=1}^{\infty}$ of partitions for a set $D \in \mathscr{D}$.

(2.4) For each point $\omega \in \Omega$, there is an integer n_ω such that, for every integer $n \geq n_\omega$, and for *every* set $C \in \mathscr{P}_n$, we have $\Phi(C)(\omega) \leq 1$.

To establish (2.4), we suppose to the contrary that each partition $\mathscr{P}_n$ contains a set C_n for which $\Phi(C_n)(\omega) > 1$. Let a pair $\{i, j\}$ of distinct positive integers be colored "red" if $C_i \cap C_j = \phi$, and let it be colored "blue" otherwise. By (2.1), we have $C_i \supseteq C_j$ for any "blue" pair $\{i, j\}$ such that $i < j$. Let $\{n_i\}_{i=1}^\infty$ be a strictly increasing sequence all of whose pairs have the same color. Suppose that the color is "red." Then the sets C_{n_i} are pairwise disjoint, and this contradicts the fact that $\Phi(D)(\omega) < \infty$. Suppose that the color is "blue." Then the C_{n_i} form a decreasing subsequence of the C_n whose intersection C is small by (2.2). It follows that $\Phi(C_{n_i})(\omega) \downarrow \Phi(C)(\omega) \leq 1$, and this contradicts the fact that $\Phi(C_{n_i})(\omega) \geq 2$ for all i. Therefore (2.4) is established.

An identical argument will also yield the following fact.

(2.5) Let m be a finite, countably additive measure on $\mathscr{D}$ such that every small set has m-measure zero. Then for every number $\varepsilon > 0$, there is an integer n_ε such that, for every integer $n \geq n_\varepsilon$, and for *every* set $C \in \mathscr{P}_n$, we have $m(C) < \varepsilon$.

We shall shortly apply this fact to the measure $m(D) = E\Phi(D)$.

3. The parameters $\mu(D)$ of the family $\Phi(D)$. If the random variable $\Phi(D)$ were Poisson, then its single parameter $\mu(D)$ would be determined by the identity

$$\mu(D) = -\ln P(\Phi(D) = 0). \tag{3.1}$$

For the moment we shall content ourselves with the observation that the right hand side of (3.1) is well defined for every set $D \in \mathscr{D}$. If it were not, then we would have $\Phi(D) \geq 1$ a.s. for at least one set $D \in \mathscr{D}$. Let $\{\mathscr{P}_n\}_{n=1}^\infty$ be a fundamental sequence of partitions for D. Then the identity $0 = P(\Phi(D) = 0) = \prod_{C\in\mathscr{P}_1} P(\Phi(C) = 0)$ shows that $\Phi(C_1) \geq 1$ a.s. for at least one set $C_1 \in \mathscr{P}_1$. We repeat this argument with C_1 in place of D and continue by induction to produce a decreasing sequence of sets $C_n \in \mathscr{P}_n$ such that $\Phi(C_n) \geq 1$ a.s. for all $n = 1, 2, \cdots$. By (2.2), the set $C = \bigcap_{n=1}^\infty C_n$ is small, and therefore $\Phi(C) = 0$ a.s. But it follows from the fact that $\Phi(C_n) \downarrow \Phi(C)$ that $\Phi(C) \geq 1$ a.s. This is a contradiction.

A glance at the complete randomness of Φ now suffices to complete the proof of the following assertion.

(3.2) The set function μ of (3.1) is well defined and constitutes a nonnegative, finitely additive measure on $\mathscr{D}$.

In fact μ is countably additive, as is easily checked, but we shall not make use of this fact.

4. The finiteness of $m(D) = E\Phi(D)$. We shall wish to apply (2.5) to the set function m defined, for all $D \in \mathscr{D}$, by $m(D) = E\Phi(D)$. It is clear that m is countably additive and that m vanishes on small sets, and so it remains to check that $m(D) < \infty$ for all $D \in \mathscr{D}$. To this end, let $\{\mathscr{P}_n\}_{n=1}^\infty$ be a fundamental sequence of partitions for D, and, for $n = 1, 2, \cdots$, define $f_n = \sum_{C\in\mathscr{P}_n} \mathscr{I}(\Phi(C) \geq 1)$, where $\mathscr{I}F$ denotes the indicator function of the set F. Then it is clear from (2.4) that $f_n \uparrow \Phi(D)$, and so it follows from the monotone convergence theorem that $Ef_n \uparrow E\Phi(D)$, or, in other words:

$$\textstyle\sum_{C\in\mathscr{P}_n} P(\Phi(C) \geq 1) \uparrow m(D). \tag{4.1}$$

By elementary calculus, we have $P(\Phi(C) \geq 1) \leq \mu(C)$, and so (3.2) and (4.1) combine to yield

$$m(D) \leq \mu(D) < \infty, \tag{4.2}$$

as desired.

The finiteness of m has one further consequence which we now note. For $n = 1, 2, \cdots$, define $g_n = \sum_{C\in\mathscr{P}_n} \mathscr{I}(\Phi(C) \geq 2) \leq \Phi(D)$. Then it is clear from (2.4) that $g_n \to 0$, and so it follows from the dominated convergence theorem that $Eg_n \to 0$, or, in other words:

$$\sum_{C\in \mathscr{P}_n} P(\Phi(C) \geq 2) \to 0. \tag{4.3}$$

This observation completes the preparation for our central argument.

5. Derivation of the Poisson distribution. If $\Phi(D)$ were Poisson, then its parameter would (also) be given by

$$\nu(D) = \frac{P(\Phi(D) = 1)}{P(\Phi(D) = 0)} = e^{\mu(D)}P(\Phi(D) = 1). \tag{5.1}$$

In view of (3.2), the complete randomness of Φ implies easily that ν is finitely additive, and we shall base an inductive argument upon the finite additivity of Φ, of μ, and of ν.

Thus, for $n = 0, 1, 2, \cdots$, and for each set $D \in \mathscr{D}$, define

$$\alpha_n(D) = e^{\mu(D)}P(\Phi(D) = n) - \nu(D)^n/n!. \tag{5.2}$$

It will be sufficient to show that $\alpha_n = 0$ for all n, since the identity $1 = \sum_{n=0}^{\infty} P(\Phi(D) = n) = e^{-\mu(D)}e^{\nu(D)}$ reveals at once that $\Phi(D)$ is Poisson with parameter $\mu(D) = \nu(D)$. Clearly we have $\alpha_0 = \alpha_1 = 0$, and so, by induction, we may assume that $n > 1$, and that $\alpha_0 = \alpha_1 = \cdots = \alpha_{n-1} = 0$. Let us first verify that α_n is finitely additive. To this end, let $D, E \in \mathscr{D}$ be disjoint. From the binomial theorem, and from the standard convolution formula for the distribution of the sum of independent random variables, we obtain the identity

$$\alpha_n(D \cup E) = e^{\mu(D)}e^{\mu(E)} \sum_{i=0}^{n} P(\Phi(D) = i)P(\Phi(E) = n - i) - \frac{1}{n!}\sum_{i=0}^{n} \binom{n}{i}\nu(D)^i\nu(E)^{n-i},$$

and this in turn is $= \alpha_n(D) + \alpha_n(E)$ because, by the induction assumption, the terms for $1 \leq i \leq n - 1$ cancel. Now let any set $D \in \mathscr{D}$ be fixed, and let $\varepsilon > 0$ be arbitrary. By (4.2), (2.5), and (4.3), there exists a (finite) partition $\mathscr{P} \subseteq \mathscr{D}$ of D such that $m(C) < \varepsilon$ for each set $C \in \mathscr{P}$, and also such that $\sum_{C\in\mathscr{P}} P(\Phi(C) \geq 2) < \varepsilon$. Since $P(\Phi(C) = 1) \leq E\Phi(C) = m(C)$, and since $P(\Phi(C) = n) \leq P(\Phi(C) \geq 2)$, it follows easily from (5.1) and (5.2) that

$$|\alpha_n(D)| \leq \sum_{C\in\mathscr{P}} |\alpha_n(C)| \leq \varepsilon(e^{\mu(D)} + m(D)e^{n\mu(D)}).$$

It is now clear from the arbitrary choice of ε that $\alpha_n(D) = 0$, and so the proof of Theorem 1.1 is complete.

6. The general case. We conclude this note with a brief discussion of how close an arbitrary completely random measure Φ comes to satisfying the hypotheses of our main theorem. The key point, due to Kingman [5, page 67], is to associate with Φ a canonical countably additive measure λ defined, for all $D \in \mathscr{D}$, by

$$\lambda(D) = -\ln E(e^{-\Phi(D)}).$$

(The finite additivity of λ is immediate from the complete randomness of Φ, while the fact that $D_n \uparrow D$ implies $\lambda(D_n) \uparrow \lambda(D)$ follows from the dominated convergence theorem, and from the continuity of ln.) We may think of λ as a kind of "control measure" for Φ, at least in the sense of the following obvious relations.

$$\begin{aligned} &\lambda(D) = 0 \quad \text{if and only if} \quad \Phi(D) = 0 \quad \text{a.s.;} \\ &\lambda(D) = \infty \quad \text{if and only if} \quad \Phi(D) = \infty \quad \text{a.s.} \end{aligned} \tag{6.1}$$

The standard partition of a set $D \in \mathscr{D}$ of finite λ-measure into at most countably many λ-atoms together with a purely nonatomic set yields a corresponding decomposition of $\Phi(D)$. If $A \in \mathscr{D}$ is a λ-atom, then it is immediate from (6.1) and from the definition of an atom [2, page 168] that, for every subset $D \in \mathscr{D}$ of A, we have either that $\Phi(D) = 0$ a.s. or that $\Phi(D) = \Phi(A)$ a.s. Thus the single random variable $X_A = \Phi(A)$ completely describes the behavior of Φ on A, and the assumption of complete randomness imposes no restrictions upon the distribution of X_A. (But let it be emphasized that the X_A's corresponding to distinct λ-atoms are mutually independent.) It is precisely for a purely nonatomic set $G \in \mathscr{D}$ that we come close to the circumstances required by Theorem 1.1, provided that Φ is integer-valued on the subsets of G.

6.2. THEOREM. *Let Φ be an integer-valued completely random measure* (*as in Theorem* 1.1), *let λ be the canonical measure associated with Φ, and let $G \in \mathscr{D}$ contain no λ-atoms. If every λ-null subset of G is small, then $\Phi(D)$ is Poisson for every subset $D \in \mathscr{D}$ of G.*

NOTE. While it is convenient to state this theorem in terms of λ, the first relation in (6.1) could be used to reformulate it in terms of Φ alone.

PROOF. It is clear from the well known special case of Liapounoff's theorem [2, Exercise (2), page 174] that there exists a countable family $\mathscr{B} \subseteq \mathscr{D}$ such that $G \subseteq \cup \mathscr{B}$, and such that, for all $x \in G$, the set $\cap \{B \in \mathscr{B} : x \in B\}$ is λ-null. (Indeed, this is perhaps most easily seen by constructing a sequence of finite partitions of G which satisfies (2.1) and (2.2) with "λ-null" in place of "small.") Since λ-null implies small in this case, the result follows.

6.3. REMARK. In [5, page 62] it seems to be assumed that λ-atoms are singletons. This assumption is of course valid when λ is a Radon measure, but it is not valid in general [1, Example 4.6, page 46]. As a consequence, Theorem 1 on page 63 of [5] could possibly require some slight modification.

REFERENCES

[1] BILLINGSLEY, P. (1979). *Probability and Measure.* Wiley, New York.
[2] HALMOS, P. R. (1950). *Measure Theory.* van Nostrand, Princeton.
[3] JAGERS, P. (1974). Aspects of random measures and point processes. In *Advances in Probability and Related Topics 3* 179–239. Marcel-Dekker, New York.
[4] KALLENBERG, O. (1976). *Random Measures.* Academic Press, London.
[5] KINGMAN, J. F. C. (1967). Completely random measures. *Pac. J. Math.* **21** 59–78.
[6] KOUMOULLIS, G. and PRIKRY, K. (1983). The Ramsey property and measurable selections. *Bull. Lond. Math. Soc.,* to appear.
[7] MATTHES, K., KERSTAN, J. and MECKE, J. (1978). *Infinitely Divisible Point Processes.* Wiley, New York.
[8] ODELL, E. (1980). Applications of Ramsey theorems to Banach space theory. In *Notes in Banach Spaces,* 379–404. University of Texas Press, Austin.
[9] PRÉKOPÀ, A. (1958). On secondary processes generated by a random point distribution of Poisson type. *Ann. Univ. Sci. Budapest. Sect. Math.* **1** 153–170.
[10] RIPLEY, B. D. (1976). Locally finite random sets: Foundations for point process theory. *Ann. Probab.* 4 983–994.

DEPARTMENT OF MATHEMATICS
MONASH UNIVERSITY
CLAYTON, VICTORIA 3168
AUSTRALIA