

NECESSARY AND SUFFICIENT CONDITIONS FOR POISSON'S DISTRIBUTION

B. O. KOOPMAN

1. **Introduction.** Poisson's law of small numbers is conventionally described as the distribution of the number of successes in a very large number of independent trials of very small individual probability of success.

The usual method of deriving Poisson's distribution is to set the individual probability of success equal to m/n in the binomial distribution of n repeated trials, and then pass to the limit as $n \rightarrow \infty$, $m = \text{constant} \geq 0$: The probability $P_n(s)$ of exactly s successes approaches the Poisson value $P(s)$,

$$(1.1) \quad P(s) = e^{-m} m^s / s! \quad (m \geq 0; s = 0, 1, 2, \dots).$$

But this method requires that at the n th stage of approximation (that is, when n trials are envisaged) the individual probabilities of success shall all be equal. This is a severe restriction, both from the theoretical point of view and from that of practical application.

In the most general case of an infinite sequence of sets of n independent trials ($n = 1, 2, 3, \dots$), we have an infinite triangular array of probabilities $p_{n,k}$ ($k = 1, \dots, n; n = 1, 2, \dots$), where $p_{n,k}$ is the probability that the k th trial in the n th set shall succeed. If $P_n(s)$ is the probability of just s success in the n th set of trials ($0 \leq s \leq n$), the question is naturally raised as to when, for each s , $P_n(s) \rightarrow P(s)$ as $n \rightarrow \infty$.

The first task of this paper is to establish simple necessary and sufficient conditions for this limit.¹ The second task is to do the same when the probabilities $p_{n,k}$ keep in a pre-assigned ratio independent of n :

$$(1.2) \quad p_{n,1} : p_{n,2} : \dots = a_1 : a_2 : \dots,$$

where the sequence $\{a_i\}$ is given. The connection of this case with divergent non-negative series satisfying a certain anti-gap condition (that is, anti-Hadamard-gap) is made apparent. The third part is devoted to the application of our methods to concrete cases. We exhibit, incidentally, examples in which the Poisson distribution is ob-

Presented to the Society, October 29, 1949; received by the editors October 10, 1949.

¹ Cf. also J. H. Curtiss, *Convergent sequences of probability distributions*, Amer. Math. Monthly (1943) pp. 97-98.

tained with such unequal $p_{n,k}$ that

$$(1.3) \quad \max_{1 \leq k \leq n} p_{n,k} / \min_{1 \leq k \leq n} p_{n,k}$$

becomes infinite as any given constant power of n . Information is supplied as to the order of approach to the Poisson distribution.

The following formulas will be needed; they are all evident consequences of simple probability reasoning.

If $m(n)$ is the expected number of successes in the set of n trials, we have

$$(1.4) \quad m(n) = \sum_{k=1}^n p_{n,k},$$

$$(1.5) \quad m(n) = \sum_{s=0}^n sP_n(s).$$

The generating function of this distribution is given by

$$(1.6) \quad \phi_n(t) = \sum_{s=0}^n P_n(s)t^s = \prod_{k=1}^n [1 + (t-1)p_{n,k}];$$

and $P_n(0)$ is easily seen to be

$$(1.7) \quad P_n(0) = \prod_{k=1}^n (1 - p_{n,k}).$$

For the Poisson distribution defined by (1.1), the "mean" or expected value of s is given by

$$(1.8) \quad m = \sum_{s=1}^{\infty} sP(s) = \sum_{s=0}^{\infty} \frac{s}{s!} e^{-m} m^s.$$

The generating function is

$$(1.9) \quad \phi(t) = \sum_{s=0}^{\infty} P(s)t^s = e^{m(t-1)}.$$

2. Necessary conditions. We now prove the following:

THEOREM I. *A necessary condition for*

$$(2.1) \quad \lim_{n \rightarrow \infty} P_n(s) = P(s)$$

is that both of the following equations hold:

$$(2.2) \quad \lim_{n \rightarrow \infty} |m(n)|_n^{\frac{1}{n}} = m,$$

$$(2.3) \quad \lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} p_{n,k} = 0.$$

First, we will show that

$$(2.4) \quad \liminf_{n \rightarrow \infty} m(n) \geq m.$$

Let $\epsilon > 0$ be arbitrarily given. Since series (1.8) is convergent, there exists an N such that

$$m < \sum_{s=0}^N sP(s) + \epsilon.$$

On the other hand, by (2.1),

$$\sum_{s=0}^N sP(s) = \lim_{n \rightarrow \infty} \sum_{s=0}^N sP_n(s).$$

But for all $n \geq N$, (1.5) shows that

$$\sum_{s=0}^N sP_n(s) \leq m(n);$$

so that, in view of the preceding inequalities,

$$\liminf_{n \rightarrow \infty} m(n) > m - \epsilon;$$

and the truth of this for arbitrary ϵ leads to (2.4).

For sufficiently large n , $p_{n,k} \neq 1$; otherwise there will be infinitely many values of n for which $P_n(0) = 0$ (cf. (1.7)), which would lead, by (2.1), to $P(0) = 0$, contrary to (1.1). We may therefore take logarithms in (1.7) and apply the remainder theorem, obtaining

$$\log P_n(0) = - \sum_{k=1}^n p_{n,k} - \sum_{k=1}^n \frac{p_{n,k}}{2(1 - \theta_{n,k} p_{n,k})^2} \quad (0 < \theta_{n,k} < 1).$$

In view of (1.4) and the fact that $(1 - \theta_{n,k} p_{n,k})^2 < 1$, this leads to

$$\log P_n(0) \leq -m(n) - \frac{1}{2} \sum_{k=1}^n p_{n,k}^2.$$

Subtracting this from $\log P(0)$ obtained from (1.1), we have

$$(2.5) \quad \begin{aligned} \log P(0) - \log P_n(0) &\geq [m(n) - m] + \frac{1}{2} \sum_{k=1}^n p_{n,k}^2 \\ &\geq [m(n) - m] + \frac{1}{2} \max_{1 \leq k \leq n} p_{n,k}^2 \geq m(n) - m. \end{aligned}$$

The third inequality contained in (2.5), taken in connection with (2.1) and (2.4), establishes (2.2). This result applied to the second inequality in (2.5), and (2.1), leads to (2.3). This proves the theorem.

3. Sufficiency of the conditions. We now show that the necessary conditions for the Poisson distribution of Theorem I are also sufficient. The first step is to prove the following:

THEOREM II. *If equations (2.2) and (2.3) are satisfied, then $\phi_n(t) \rightarrow \phi(t)$ uniformly in any given circle $|t-1| \leq r$ of the plane of the complex variable t . Moreover, for that circle,*

$$(3.1) \quad \phi_n(t) = \phi(t) + O[m(n) - m] + O\left(\max_{1 \leq k \leq n} p_{n,k}\right).$$

Here the order symbol O denotes *uniform* order: If $\{c_n\}$ is an infinite sequence of constants and $\{\psi_n\}$ a sequence of functions defined on a common domain D , $\psi_n = O(c_n)$ signifies the existence of a constant $K (=K_D)$ such that, for all sufficiently large n , $|\psi_n| \leq K|c_n|$ holds throughout D . In (3.1), $D: (|t-1| \leq r)$.

In view of (2.3), for all sufficiently large n

$$(3.2) \quad |(t-1)p_{n,k}| \leq r \max_{1 \leq k \leq n} p_{n,k} \leq \frac{1}{2}.$$

This restriction upon n shall be maintained throughout the proof.

Let $\log(1+u)$ denote, for all $|u| \leq 1/2$, that determination of the logarithm whose angle is between $\pm\pi$ (actually, it will be between $\pm\pi/6$). An elementary integration shows that

$$(3.3) \quad \log(1+u) = u - u^2 \int_0^1 \frac{(1-v)dv}{(1+uv)^2},$$

where the path of integration is along the real axis from 0 to 1.

Taking logarithms in (1.6) and applying (3.3) (with $u = (t-1)p_{n,k}$) and (1.4), we obtain

$$(3.4) \quad \log \phi_n(t) = (t-1)m(n) - w_n(t),$$

where

$$w_n(t) = (t-1)^2 \sum_{k=1}^n p_{n,k}^2 \int_0^1 \frac{(1-v)dv}{[1+(t-1)p_{n,k}v]^2}.$$

Using (3.2) we find

$$\begin{aligned}
 |w_n(t)| &\leq |t-1|^2 \sum_{k=1}^n p_{n,k}^2 \int_0^1 \frac{(1-v)dv}{[1-|(t-1)p_{n,k}|v]^2} \\
 &\leq r^2 \sum_{k=1}^n p_{n,k}^2 \int_0^1 \frac{(1-v)dv}{(1-1/2)^2} \\
 &= 2r^2 \sum_{k=1}^n p_{n,k}^2;
 \end{aligned}$$

whence, finally,

$$\begin{aligned}
 (3.5) \quad |w_n(t)| &\leq 2r^2 \left(\max_{1 \leq k \leq n} p_{n,k} \right) \sum_{k=1}^n p_{n,k} \\
 &= 2r^2 m(n) \max_{1 \leq k \leq n} p_{n,k}.
 \end{aligned}$$

On writing (3.4) in the form

$$\log \phi_n(t) = (t-1)m + (t-1)[m(n) - m] - w_n(t),$$

and noting that in virtue of (2.2), (2.3), (3.5),

$$(t-1)[m(n) - n] = O[m(n) - m], \quad w_n(t) = O\left(\max_{1 \leq k \leq n} p_{n,k}\right),$$

equation (3.1), and therewith our theorem, become evident.

THEOREM III. *Equations (2.2) and (2.3) are sufficient to ensure that (2.1) hold uniformly in s ; and, moreover, that*

$$(3.6) \quad P_n(s) = P(s) + O[m(n) - m] + O\left(\max_{1 \leq k \leq n} p_{n,k}\right).$$

Here O denotes uniform order for s ranging over the domain $D: (0 \leq s \leq \infty)$.

We apply Theorem II with any fixed $r \geq 2$ and denote by C the circle $|t-1|=r$. Expressing $P_n(s)$ and $P(s)$ as contour integrals about C by the Cauchy integral formula, and applying (3.1), we obtain the conclusion (3.6) immediately.

Thus equations (2.2) and (2.3) are necessary and sufficient for the Poisson distribution, (2.1); they shall be called the *Poisson conditions*.

We remark that when $m > 0$, a slight modification in the approximating process leads to a simpler asymptotic relation than (3.6): If $n_{n,k}$ satisfy the Poisson conditions with $m > 0$, so will

$$(3.7) \quad p'_{n,k} = \frac{m}{m(n)} p_{n,k},$$

once n is sufficiently large. And since the corresponding $m'(n) = m$, we obtain the same Poisson distribution, only with (3.6) replaced by

$$(3.8) \quad P'_n(s) = P(s) + O(\max_{1 \leq k \leq n} p'_{n,k}).$$

4. The case of proportional probabilities. Many practical applications suggest the consideration of the array $\{p_{n,k}\}$ which satisfies, in addition to the Poisson conditions, the further *proportionality condition* (1.2). Algebraically, this means that for all i, j from 1 to n , and for all n ,

$$(4.1) \quad a_i p_{n,j} = a_j p_{n,i}.$$

Evidently we can assume that the a 's are non-negative and not all zero.

There is no essential loss in generality in assuming $a_1 \neq 0$. For there is always a first a , call it a_{N+1} , which is not zero. Applying (4.1) with $i = N+1$ and $j \leq N$, we see that for all $n > N$,

$$p_{n,1} = p_{n,2} = \dots = p_{n,N} = 0.$$

Discard the first N sets of trials. In each of the infinitude of remaining sets, ignore the first N trials: All these have probability zero. In other words, consider the infinite triangular array $\{p'_{n,k}\}$, where

$$p'_{n,k} = p_{N+n,N+k} \quad (1 \leq k \leq n; n = 1, 2, \dots).$$

This is the sub-array of the original array $\{p_{n,k}\}$, lying to the right of the vertical line drawn through the column of quantities $p_{n,N}$ ($n = N, N+1, \dots$). On setting $a'_k = a_{N+k}$, the proportionality conditions (4.1) (p and a being accented) are still valid; but now $a'_1 \neq 0$. Furthermore, if $\{p_{n,k}\}$ satisfies the Poisson conditions, so will $\{p'_{n,k}\}$, and *with the same mean m .*

THEOREM IV. *Let $\{p_{k,k}\}$ satisfy the Poisson conditions with $m > 0$, and also the proportionality conditions (4.1) with $a_1 > 0$. Then the series $a_1 + a_2 + \dots$ diverges and satisfies the "anti-gap condition"*

$$(4.2) \quad \lim_{n \rightarrow \infty} \frac{a_n}{s_n} = 0, \quad s_n = a_1 + \dots + a_n.$$

Furthermore, for all sufficiently large n ,

$$(4.3) \quad p_{n,k} = m(n)a_k/s_n \quad (1 \leq k \leq n).$$

Suppose that, for a particular n , $p_{n,1} = 0$. Applying (4.1) with $i = 1$ and $j > 1$, we see that $p_{n,j} = 0$; so that $m(n) = 0$. This cannot occur for infinitely many values of n , since $m \neq 0$. Therefore, for all sufficiently large n , $m(n) \neq 0$ and $p_{n,1} \neq 0$. We now confine our attention to such n .

Set $i = 1, j = k$, in (4.1) and solve for a_k :

$$(4.4) \quad a_k = a_1 p_{n,k} / p_{n,1} \quad (k = 1, 2, \dots, n);$$

whence

$$(4.5) \quad s_n = a_1 m(n) / p_{n,1}.$$

It follows from the Poisson conditions ($m > 0$) that $s_n \rightarrow \infty$. Further, (4.4) and (4.5) show that

$$(4.6) \quad \frac{a_k}{s_n} = \frac{p_{n,k}}{m(n)} \quad (1 \leq k \leq n),$$

from which the Poisson conditions lead to (4.2). And (4.3) follows from (4.6). This completes the proof.

THEOREM V. *Let $a_1 + a_2 + \dots$ be any divergent non-negative term series satisfying the anti-gap condition (4.2); and let $m(n)$ be any non-negative sequence converging to any $m > 0$. Then if $p_{n,k}$ are defined by (4.3), they will satisfy the Poisson and proportionality conditions (4.1).*

Equations (4.1) and (2.2) being evident, we have but to establish (2.3). Let $\mu(n)$ be the subscript i of that a_i which is the (first) maximum in the set $\{a_1, \dots, a_n\}$. Clearly

$$\max_{1 \leq k \leq n} p_{n,k} = \frac{m(n)a_{\mu(n)}}{s_n}.$$

Now $\mu(n)$, which is nondecreasing as n increases, either remains bounded as $n \rightarrow \infty$ (in which case (2.3) is an evident consequence of $s_n \rightarrow \infty$) or else becomes infinite. In the latter case we have

$$\max_{1 \leq k \leq n} p_{n,k} = \frac{m(n)a_{\mu(n)}}{s_{\mu(n)}} \cdot \frac{s_{\mu(n)}}{s_n} \leq \frac{m(n)a_{\mu(n)}}{s_{\mu(n)}};$$

(2.3) is now a consequence of the anti-gap condition.

An example, of practical significance, of the Poisson and proportionality conditions is furnished by the sequence $\{c_i\}$ ($c_i \geq 0$) having a limiting arithmetic mean m of its first n terms, as $n \rightarrow \infty$. If $p_{n,k} = c_k/n$, the Poisson and proportionality conditions ((4.1) with $a_k = c_k$) may be seen to apply, leading to (2.1). But we may generalize this type of example by extending the notion of arithmetic mean.

Let us call any sequence $\{\omega_i\}$ an *averaging sequence* if it has the three properties:

$$(4.7) \quad 0 < \omega_1 \leq \omega_2 \leq \dots$$

$$(4.8) \quad \lim_{n \rightarrow \infty} \omega_n = \infty,$$

$$(4.9) \quad \lim_{n \rightarrow \infty} \frac{\omega_{n+1}}{\omega_n} = 1;$$

and call a sequence $\{c_i\}$ ($c_i \geq 0$) ω -*averagable* if

$$(4.10) \quad \lim_{n \rightarrow \infty} (c_1 + \dots + c_n)/\omega_n = m.$$

Under these conditions, and if $m > 0$, it will follow that the array

$$(4.11) \quad p_{n,k} = c_k/\omega_n$$

satisfies the hypothesis of Theorem V.

Equations (2.2) and (4.1) being obvious, we have but to verify (2.3). First note that, if $s_n = c_1 + \dots + c_n$,

$$\frac{c_n}{\omega_n} = \left(\frac{s_n}{\omega_n} - \frac{s_{n-1}}{\omega_{n-1}} \right) + \left(1 - \frac{\omega_{n-1}}{\omega_n} \right) \frac{s_{n-1}}{\omega_{n-1}}.$$

Then (4.9) and (4.10) show that $c_n/\omega_n \rightarrow 0$. Let $\mu(n)$ be the subscript i of the (first) maximum c_i in $\{c_1, \dots, c_n\}$; clearly

$$\max_{1 \leq k \leq n} p_{k,n} = \frac{c_{\mu(n)}}{\omega_n} = \frac{c_{\mu(n)}}{\omega_{\mu(n)}} \frac{\omega_{\mu(n)}}{\omega_n}.$$

From this it follows, using (4.8) if $\mu(n)$ is bounded, and (4.7) and $c_n/\omega_n \rightarrow 0$ when $\mu(n) \rightarrow \infty$, that (2.3) is true in all cases.

Examples of this theorem occur when $\omega_n = n^r$ ($r > 0$), $\omega_n = \sum_{h=1}^n h^{-r}$ ($0 < r \leq 1$), $\omega_n = \log n$, and so on. But if $\omega_n = r^n$ ($r \neq 0$), the hypothesis fails.

5. Illustrations. We conclude this note with three examples of the present method, the first showing its quantitative flexibility, the second and third being of practical significance; the third will also illustrate a generalization of the proportionality conditions.

EXAMPLE 1. Let $\{b_n\}$ be a sequence bounded as follows:

$$(5.1) \quad 0 < \alpha \leq b_n \leq \beta < \infty \quad (n = 1, 2, \dots).$$

On setting

$$(5.2) \quad a_n = b_n/n^r \quad (0 \leq r \leq 1)$$

and taking $p_{n,k} = ma_k/s_n$ ($s_n = a_1 + \dots + a_n$), we easily see that the hypothesis of Theorem V is valid. This leads to (2.1) and (4.1); and the ratio of (1.3) is between $(\alpha/\beta)n^r$ and $(\beta/\alpha)n^r$. Referring to (3.8), the use of obvious inequalities shows that

$$\begin{aligned}
 P_n(s) &= P(s) + O\left(\frac{1}{n^{1-r}}\right) && (0 \leq r < 1); \\
 &= P(s) + O\left(\frac{1}{\log n}\right) && (r = 1).
 \end{aligned}$$

Similarly, if we replace (5.2) by

$$(5.3) \qquad a_n = b_n n^r \qquad (r \geq 0),$$

$p_{n,k} = ma_k/s_n$ as before. The ratio (1.3) is again between $(\alpha/\beta)n^r$ and $(\beta/\alpha)n^r$, while (3.8) leads to

$$P_n(s) = P(s) + O\left(\frac{1}{n}\right).$$

EXAMPLE 2. Let an object, specified by a coordinate x , move with x going from large negative to large positive values, and, at each crossing of an integral value, be exposed to a chance of an event which we shall call a "hit." The probabilities of various numbers s of hits is required. It is assumed that the probabilities of a hit at the various integral points are proportional to given numbers, and that the total expected number of hits has a given value m . One is interested in the case where all the risks of hit are very small but the number of danger points crossed is very large: mathematically, in a limiting process.

In practice, the situation can almost always be put in the following form: A function $f(x)$ is given for all x ; it is continuous and non-negative everywhere, and approaches zero as $x \rightarrow \pm \infty$. It is bounded and has a finite number of relative maxima on $(-\infty, \infty)$. Finally, its integrals over $(0, \infty)$ and $(0, -\infty)$ diverge. The probabilities of hits at the integral points $x = k$ are proportional to $f(k)$.

Let l be any positive integer, set $n = 2l + 1$, and write

$$k = -l, -l + 1, \dots, -1, 0, 1, \dots, l - 1, l.$$

If, then, we take for the probability of a hit at $x = k$

$$p_{n,k} = m \frac{f(k)}{\sum_{i=-l}^l f(i)},$$

the conditions of our problem will be satisfied at the n th approximation. What occurs when $n \rightarrow \infty$ (that is, $l \rightarrow \infty$)?

Since $f(x)$ has but a finite number of relative maxima, a comparison of areas shows that as $l \rightarrow \infty$,

$$\sum_{i=-l}^l f(i) \rightarrow \infty.$$

On the other hand, since $f(x) \leq M$ (boundedness),

$$\max_{-l \leq k \leq l} p_{n,k} \leq \frac{M}{\sum_{i=-l}^l f(i)} \rightarrow 0$$

as $l \rightarrow \infty$. Thus the Poisson conditions are satisfied—with this immaterial change of notation for k , and the equally immaterial restriction of n to being odd. The Poisson distribution is obtained, the limiting probability of just s hits as the object moves from $-\infty$ to $+\infty$ being $P(s)$.

EXAMPLE 3. Another exposure problem is the following: The region of exposure is the fixed interval $a \leq x \leq b$, which the object crosses once, from left to right. On this interval a continuous non-negative function $f(x)$ is given. At the n th stage of approximation, risk of hit occurs at the n points $x = a + k(b-a)/n$ ($k = 1, 2, \dots, n$). The probabilities of hit are

$$p_{n,k} = \frac{1}{n} f(a + k(b-a)/n).$$

We have (setting $x_k = a + k(b-a)/n$ and $\Delta x = (b-a)/n$)

$$\begin{aligned} m(n) &= \frac{1}{n} \sum_{k=1}^n f(a + k(b-a)/n) \\ &= \frac{1}{b-a} \sum_{k=1}^n f(x_k) \Delta x \rightarrow \frac{1}{b-a} \int_a^b f(x) dx = m; \end{aligned}$$

and (since $f(x)$ is bounded on $a \leq x \leq b$)

$$\max_{1 \leq k \leq n} p_{n,k} \leq \frac{M}{n} \rightarrow 0.$$

Thus the Poisson distribution of mean m is obtained. But the probability ratios vary with n . They will, however, approach limiting values (cf. (5.1)). For simplicity, suppose that $f(x) > 0$ ($a \leq x \leq b$).

Then

$$\frac{p(n, k+1)}{p(n, k)} = \frac{f(x_k + \Delta x)}{f(x_k)} = 1 + \frac{f(x_k + \Delta x) - f(x_k)}{f(x_k)}$$

and, by the uniform continuity of $f(x)$ on the closed interval $a \leq x \leq b$, we have $f(x_k + \Delta x) - f(x_k) \rightarrow 0$ and $n \rightarrow \infty$. And since our conditions insure that $f(x)$ is bounded away from zero, $p(n, k+1)/p(n, k) \rightarrow 1$: The limiting ratios are all unity.

This example leads one to seek a broadening of the proportionality conditions, by requiring that (4.1) apply only in a limiting sense. A formulation appropriate to probability is the following:

Let (4.1) be replaced by the requirement that any array of probabilities $p'_{n,k}$ shall exist which itself satisfies (4.1), and is related to the original $p_{n,k}$ by the equations

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} |p_{n,k} - p'_{n,k}| = 0, \quad \lim_{n \rightarrow \infty} [m(n) - m'(n)] = 0,$$

where $m'(n)$ is the expected value for the accented set. Then it is seen at once that if $p_{n,k}$ satisfy (2.2), (2.3), so do $p'_{n,k}$, and with the same m ; the latter thus lead to the same Poisson distribution as the former. And, of course, Theorems IV, V apply to $a_1 + a_2 + \dots$.

In Example 3, we are thus lead to setting $a_1 = a_2 = \dots = 1$ and $p'_{n,k} = m/n$, where m is still the mean of $f(x)$ over (a, b) . The function $f_1(x) = m$ in $(a \leq x \leq b)$, equal to 0 outside, is related to $p'_{n,k}$ in just the same way as the original $f(x)$ was related to $p_{n,k}$; $f_1(x)$ may be described as a sort of "effective risk function" replacing $f(x)$.

Similar considerations in the case where $f(x)$ is zero in certain sub-intervals of (a, b) lead to an $\{a_n\}$ sequence of zeros and ones, and to an effective risk function $f_1(x)$ equal to m when $f(x) > 0$ and 0 when $f(x) = 0$.

It may be remarked that in Examples 2, 3, the x -axis may be the time axis and the danger points, epochs.