A THEORY OF THE SPATIAL DISTRIBUTION OF GALAXIES*

J. Neyman and E. L. Scott
Statistical Laboratory, University of California
Received February 18, 1952

ABSTRACT

A theory of the spatial distribution of galaxies is built, based on the following four main assumptions: (i) galaxies occur only in clusters; (ii) the number of galaxies varies from cluster to cluster, subject to a probabilistic law; (iii) the distribution of galaxies within a cluster is also subject to a probabilistic law; and (iv) the distribution of cluster centers in space is subject to a probabilistic law described as quasi-uniform. The main result obtained is the joint probability generating function $G_{N_1, N_2}(t_1, t_2)$ of numbers N_1 and N_2 of galaxies visible on photographs from two arbitrarily placed regions ω_1 and ω_2, taken with fixed limiting magnitudes m_1 and m_2, respectively. The theory ignores the possibility of light-absorbing clouds. The function $G_{N_1, N_2}(t_1, t_2)$ is expressed in terms of four functions left unspecified, which govern the details of the structure contemplated. Methods are indicated whereby approximations to these functions can be obtained and whereby the general validity of the hypotheses can be tested.

I. INTRODUCTION

The distribution of the numbers of galaxies as revealed by the data of Hubble,[1] Shapley,[2] and Shane[3] does not conform with the Poisson law and indicates the presence of a factor causing "contagion." In order to explain this phenomenon, two kinds of hypotheses are discussed in the literature. Hypotheses of one kind postulate that not only the apparent but also the actual spatial distribution of galaxies is clustered.[4] Indeed, some clusters of galaxies appear to be identified and are studied per se.[5] The hypotheses of the alternative type tend to explain the apparent clustering of galaxies by the effects of extinction of their light by interstellar clouds.[6] Undoubtedly both factors play a role.

The purpose of the present paper is to study the implications of a probabilistic model of spatial clustering of galaxies, ignoring the possibility of extinction by clouds. It is hoped that this model will fit the data relating to regions of the sky not affected by clouds. However, whether it does or not, the evaluation of the effects of clustering in space and the comparison with observations are likely to be helpful in understanding the machinery of the phenomena studied.

The probabilistic model considered is based on several postulates which we classify under two headings: postulates essential to the model, or structural postulates, on the one hand, and secondary postulates, on the other. All the postulates are described below in precise mathematical form. Here, however, it seems useful to characterize them in a manner which is less precise but which is likely to make a stronger appeal to the intuition.

The postulates labeled "structural" or "essential" are those reflecting the major properties of the universe contemplated. These are:

a) Galaxies occur only in clusters.

b) The number of galaxies varies from one cluster to another in a manner subject to a definite probabilistic law, the same for all clusters.

* This paper was prepared with the partial support of the Office of Naval Research

1 *Mt. W. Contr.*, No. 485; *Ap. J.*, **79**, 9, 1934; *Mt. W. Contr.*, No. 557; *Ap. J.*, **84**, 517, 1936.

2 *Harvard Ann.*, Vol. **88**, No. 2, 1932; Vol. **105**, No. 8, 1937; Vol. **106**, No. 1, 1938.

3 In preparation; see also *Proc. Amer. Phil. Soc.*, **94**, 13, 1950.

4 C. V. L. Charlier, *Ark. mat. astr. fys.*, Vol. **16**, No. 22, 1921; *Medd. Lunds Obs.*, Ser. II, No. 128, 1950.

5 E. Hubble and M. Humason, *Mt. W. Contr.*, No. 427, 1931; *Ap. J.*, **74**, 43, 1931; H. Shapley, *Proc. Nat. Acad. Sci.*, **19**, 591, 1933.

6 V. A. Ambartzumian, *Comm. Burakan Obs.*, issue VI, 1951.

c) The distribution of galaxies within a cluster is random and is subject to a probabilistic law which also is the same for all clusters.

d) The distribution of cluster centers in space is random and what we propose to call "quasi-uniform." To explain this term, we shall consider two cubes, k_1 and k_2, of equal volume arbitrarily placed in space but nonoverlapping. A distribution will be called "quasi-uniform" if it satisfies the following two conditions: (i) Whatever the integer $n \geqq 0$, the probability that there will be exactly n cluster centers in k_1 is always equal to the probability that there will be exactly n cluster centers in k_2. We emphasize that this equality persists regardless of the location of the two regions k_1 and k_2. (ii) The presence in k_1 of any particular number of cluster centers does not influence the probability that k_2 will contain exactly n cluster centers ($n = 0, 1, 2, \ldots$), and this is regardless of whether k_1 and k_2 are near by or distant, provided that they do not overlap.

e) Whatever may be the motions of the cluster centers and of the galaxies within the clusters, they can be neglected.

The last two postulates require some comments. Postulate *d* characterizes what we call "quasi-uniformity of a distribution." Ordinarily, when one speaks of statistical uniformity in space, one has in mind a distribution of points obtainable from a particular machinery of distributing these points. This machinery consists in dividing a bounded portion $\mathfrak{V}$ of the space into a finite but very large number N of elementary volumes $\mathfrak{v}$ of equal size. Then, in order to place a point P_1 in $\mathfrak{V}$, we select one of the elements $\mathfrak{v}$, with equal chance $1/N$ for each element, and place P_1 in the element selected. Once the point P_1 is placed, we proceed to place the next point, P_2, in exactly the same manner and without regard to the position assumed by P_1, etc. As is well known, if N and the number of points $P_1, P_2, \ldots$, are both large, then the distribution resulting from the above operations will conform with the law of Poisson: whatever be the volume $\mathfrak{V}^*$ partial to $\mathfrak{V}$, the probability that it will contain exactly n points P is given by the formula

$$e^{-\lambda} \frac{\lambda^n}{n!}, \tag{1}$$

where λ is a suitable constant. This machinery of distributing points P in the volume $\mathfrak{V}$ may be called the "Poisson machinery."

It will be seen that the Poisson machinery of distributing points possesses the characteristic of quasi-uniformity which we described. We discuss this point in some detail for the reason that the converse statement is not true, so that a quasi-uniform distribution of points need not necessarily be the Poisson distribution. In order to illustrate this circumstance, consider the following machinery of distributing points $P_1, P_2, \ldots$, in the volume $\mathfrak{V}$. As formerly, this volume $\mathfrak{V}$ is divided into a large number N of elements $\mathfrak{v}$ of equal size. Then a die is thrown, and the number d of dots on the upper face is noted. Further, out of the N elementary volumes, one is selected, with the probability of any particular selection equal to $1/N$. Then the first d points, $P_1, P_2, \ldots, P_d$, are put into the volume element selected. This process is then repeated many times. The machinery just described will satisfy postulate *d*. However, it is obviously not the Poisson machinery. It, and many similar machineries, produce what is known as *contagious* distributions of points—this in spite of the characteristic of quasi-uniformity.

The purpose of formulating postulate *d* in this particular form is to cover the possibility that the cluster centers are distributed in space independently of each other (that is, Poisson-wise) and also the possibility that they themselves are clustered. In the latter case the presence of a cluster center in a given volume would increase the probability that the volume includes some additional cluster centers.

Now we must comment on postulate *e*. Although not explicitly stated, postulates *a*–*d* refer to a particular moment of time, for example, to the moment of taking photographs. The galaxies photographed are generally at different distances from the observer and, owing to the time spent by light in traveling to the observer, the positions and the apparent magnitudes of the individual galaxies photographed refer to different moments. Should the galaxies and/or the cluster centers have substantial velocities, then the distribution of their positions at these different times need not conform with postulates *a–d*, even though the distribution of simultaneous positions does. The purpose of postulate *e*, then, is to eliminate the consideration of this possibility. An alternative postulate might be one reflecting the hypothesis of an expanding universe.

The foregoing structural postulates are used below to obtain general forms of the distributions of galaxies on the photographic plate. In order to obtain these distributions numerically, it is necessary to adopt further hypotheses specializing the distributions mentioned in the structural postulates. These additional hypotheses are described as "secondary" postulates. They can be formulated in many different ways. Our attitude toward any particular function selected is that toward an interpolation formula. The final choice will depend upon comparison with observations. In the meantime, our preference is for as few parameters as possible combined with flexibility and simplicity in the formulae.[7]

II. PRELIMINARY REMARKS

Under this heading we combine certain known formulae of the calculus of probability which will be used frequently.

1. *Fundamental formula on conditional expectations.*—In the following we shall use frequently the fundamental formula connecting conditional and absolute expectations. Let X and Y be any two random variables. Let $E(Y|X = x)$ denote the conditional expectation of Y, given that the random variable X assumed a specified value x. Further, let $E(Y)$ denote the unconditional or absolute expectation of Y. We assume that both these expectations exist. Starting with $E(Y|X = x)$, we define a new random variable $E(Y|X)$ as follows: Whenever X assumes any value x, then $E(Y|X)$ assumes the value $E(Y|X = x)$. The fundamental formula mentioned is, then,

$$E(Y) = E\{E(Y|X)\}. \tag{2}$$

2. *Probability generating function.*—Let $X_1, X_2, \ldots, X_s$ be a set of s random variables, all capable of assuming only nonnegative integer values $0, 1, 2, \ldots$. Further, let $t_1, t_2, \ldots, t_s$ be arbitrary nonnegative numbers not exceeding unity. Then the probability generating function of $X_1, X_2, \ldots, X_s$ is defined as the expectation of the product $t_1^{X_1} t_2^{X_2} \ldots t_s^{X_s}$. The generic notation for the probability generating function is

$$\begin{aligned} G_{X_1, X_2, \ldots, X_s}(t_1, t_2, \ldots, t_s) &= E\Big[\prod_{i=1}^{s} t_i^{X_i}\Big] \\ &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\cdots\sum_{n_s=0}^{\infty}\prod_{i=1}^{s} t_i^{n_i} P\{(X_1 = n_1)(X_2 = n_2)\ldots(X_s = n_s)\}, \end{aligned} \tag{3}$$

where $P\{(X_1 = n_1)(X_2 = n_2) \ldots (X_s = n_s)\}$ denotes the probability that simultaneously $X_1 = n_1, X_2 = n_2, \ldots, X_s = n_s$ and where the multiple series in the right-hand side is uniformly convergent for $0 \leqq t_i \leqq 1, i = 1, 2, \ldots, s$.

A probability generating function is differentiable indefinitely with respect to all its arguments for $0 \leqq t_i < 1, i = 1, 2, \ldots, s$. The partial derivatives of the probability generating function are connected with the probabilities $P\{(X_1 = n_1)(X_2 = n_2) \ldots$

[7] The results obtained in connection with certain secondary postulates will be published in a subsequent paper by the authors and C. D. Shane.

$(X_s = n_s)\}$ and also with the moments and product moments of the variables $X_1, X_2, \ldots, X_s$. In fact, formula (3) implies that

$$P\{(X_1=n_1)(X_2=n_2)\ldots(X_s=n_s)\} = \frac{\partial^{n_1+n_2+\ldots+n_s}G}{\partial t_1^{n_1}\partial t_2^{n_2}\ldots\partial t_s^{n_s}}\bigg|_{t_1=t_2=\ldots=t_s=0}\prod_{i=1}^{s}\frac{1}{n_i!}. \quad (4)$$

Also, if the derivative of the right-hand side of expression (4) exists at $t_1 = t_2 = \ldots = t_s = 1$, then

$$\frac{\partial^{n_1+n_2+\ldots+n_s}G}{\partial t_1^{n_1}\partial t_2^{n_2}\ldots\partial t_s^{n_s}}\bigg|_{t_1=t_2=\ldots=t_s=1} = E\left\{\prod_{i=1}^{s}\frac{X_i!}{(X_i-n_i)!}\right\}. \quad (5)$$

Because of these properties, the study of distributions of random variables capable of assuming only nonnegative integer values is reducible to the study of the corresponding probability generating functions. This method was introduced by Laplace. A modern presentation referring to single variate distributions is given in a recent book by Feller.[8]

III. GENERAL STOCHASTIC MODEL OF THE SPATIAL DISTRIBUTION OF GALAXIES

In this section we restate in mathematical form the several structural postulates which underlie the stochastic model of the universe studied in this paper. However, it is convenient to change the order in which the postulates are listed.

POSTULATE 1.—*To every region* R *with volume* $\mathfrak{V}$ *there corresponds a random variable* γ(R) *representing the number of cluster centers falling within* R. *The distribution of* γ(R) *depends on the volume* $\mathfrak{V}$ *only*. It does not depend either on the shape of the region R or on its location.

The probability generating function of $\gamma(R)$ will be denoted $G_\gamma(t|\mathfrak{V})$.

POSTULATE 2.—*Whatever be the nonoverlapping regions* $R_1, R_2, R_3, \ldots$, *the corresponding random variables* $\gamma(R_j)$, $j = 1, 2, \ldots$, *are completely independent.*

Postulates 1 and 2 deal with the probability distribution of the *number* of cluster centers within any given region R in space. In the following it will be necessary to consider as random the position of a single cluster center, given that it is located in R. This appears to require a special postulate (postulate 3) to the general effect that, as far as their positions in R are concerned, all clusters are, in a sense, equivalent. Then postulates 1, 2, and 3 combine to imply that the position of any given cluster is random. This point is discussed in some detail in Section IV.

POSTULATE 3.—*Let* R *be an arbitrary region of positive volume* $\mathfrak{V} > 0$ *and* R_1 *an arbitrary part of* R. *If* R *is known to include exactly* $n > 1$ *cluster centers numbered* $C_1, C_2, \ldots, C_n$, *and if* $a_1 < a_2 < \ldots < a_m$ *is an arbitrary combination of* $m < n$ *numbers selected out of* $1, 2, \ldots, n$, *then the probability that* R_1 *will contain exactly* m *cluster centers and that these cluster centers will be* $C_{a_1}, C_{a_2}, \ldots, C_{a_m}$ *is independent of the combination* $a_1, a_2, \ldots, a_m$ *and is equal to the conditional probability that* $\gamma(R_1) = m$, *given that* $\gamma(R) = n$, *divided by the number of combinations of* m *objects out of the given* n.

We shall use the symbol $C_i \in R_1$ to denote that the cluster center C_i is included in R_1. Also the notation[9]

$$\prod_{i=1}^{m}(C_{a_i}\in R_1) = (C_{a_1}\in R_1)(C_{a_2}\in R_1)\ldots(C_{a_m}\in R_1) \quad 6)$$

[8] *Probability Theory and Its Applications* (New York: John Wiley & Sons, Inc., 1950).

[9] This notation is very convenient and was used in formula (3). The reader unaccustomed to it may wish to consult J. Neyman, *First Course in Probability and Statistics* (New York: Henry Holt & Co., 1950).

will denote the logical product of the m propositions $C_{a_i} \in R_1$ for $i = 1, 2, \ldots, m$. The meaning of this term is: All the cluster centers $C_{a_1}, C_{a_2}, \ldots, C_{a_m}$ are included in R_1.

With this notation, postulate 3 may be expressed by the formula

$$P\left\{\left[\gamma(R_1) = m\right] \prod_{i=1}^{m} (C_{a_i} \in R_1) \mid \gamma(R) = n\right\} = \frac{m!\,(n-m)!}{n!} P\{\gamma(R_1) = m \mid \gamma(R) = n\}. \quad (7)$$

Postulate 4.—*To every cluster center there corresponds a random variable ν, representing the number of galaxies belonging to this cluster. The distribution of ν is the same for all clusters.*

The probability generating function of ν will be denoted by $G_\nu(t)$.

Postulate 5.—*The random variables $\nu_1, \nu_2, \ldots, \nu_k, \ldots$, representing the numbers of galaxies in different clusters, are completely independent.*

Further below we shall need symbols to denote the co-ordinates of a cluster center and those of a galaxy, which we shall treat as a point. In cases where a *random cluster center* is considered, its co-ordinates will be treated as random variables and denoted by the capital letters U, V, W. The particular values assumed by these variables will be denoted by the corresponding lower-case letters u, v, w. Similarly, the random variables representing a *random galaxy* will be denoted by the capitals X, Y, Z, and the particular values of the variables by x, y, z. The co-ordinate axes will be assumed orthogonal, with their origin at the observer and with arbitrary directions.

Postulate 6.—*Given that the center of a cluster is at a specified point* (u, v, w), *the position of every galaxy belonging to this cluster is random and the probability density of its co-ordinates* X, Y, Z *is represented by a function* f (η) *depending only on the distance*

$$\eta = \{(x-u)^2 + (y-v)^2 + (z-w)^2\}^{1/2} \quad (8)$$

between the cluster center and the position of the galaxy. The function f(η) *is the same for all clusters and is continuous for all values of η.*

Postulate 7.—*Whatever be the galaxies* $g_1, g_2, \ldots, g_m, \ldots$, *whether belonging to the same cluster or not, the triplets of random variables* (X_m, Y_m, Z_m) *representing their co-ordinates,* $m = 1, 2, \ldots$, *are completely independent.*

The foregoing seven postulates determine what is considered to be the structure of the spatial distribution of galaxies. As such, of course, this distribution is unobservable. The following postulate 8 is meant to establish a link between the unobservable distribution of galaxies in space and what can be observed on photographic plates. It is convenient to make this link in two steps. First we consider an idealized photographic plate which, for any given observational setup, has a fixed limiting apparent magnitude, constant for the entire plate. If the apparent magnitude of a galaxy is less than this limiting magnitude, it will sometimes be convenient to say that this galaxy is "visible" in the observational setup considered. The relation between the numbers of galaxies visible on an idealized plate and the counts on actual plates is dealt with in a later paper by the authors jointly with C. D. Shane, in which empirical data are discussed.

Postulate 8.—*Given an observational setup, to every galaxy with co-ordinates* (x, y, z) *there corresponds a probability $\theta(\xi)$ that the apparent magnitude of this galaxy will be less than the limiting apparent magnitude of the idealized photographic plate. This probability depends only upon the distance*

$$\xi = \{x^2 + y^2 + z^2\}^{1/2} \quad (9$$

to the galaxy and is continuous for all values of ξ.

It will be seen that, in order to obtain numerically the characteristics of the distribution of galaxies visible on photographic plates, many details of the foregoing general postulates will have to be specialized. Thus, for example, the probability $\theta(\xi)$ will have to be specialized in accordance with the limiting apparent magnitude of the idealized plate, with the assumed form of the distribution of absolute magnitudes of extragalactic nebulae, and with the effect of red shift. All these details are independent of the structural postulates. It is interesting, however, that postulates 1 and 2 do establish a limitation on the probability generating function $G_\gamma(t|\mathfrak{V})$ of the number of cluster centers falling in a region of volume $\mathfrak{V}$.

IV. IMPLICATIONS OF POSTULATES 1 AND 2

In this section we discuss briefly the general form of the probability generating function $G_\gamma(t|\mathfrak{V})$ as implied by the two postulates 1 and 2. Let R_1 and R_2 be two nonoverlapping regions of finite volumes $\mathfrak{V}_1$ and $\mathfrak{V}_2$, and let $R_1 \cup R_2$ be the union of R_1 and R_2. Finally, let γ_1, γ_2, and γ_3 stand for the numbers of cluster centers in R_1, R_2 and $R_1 \cup R_2$, respectively. Obviously, $\gamma_3 = \gamma_1 + \gamma_2$. Because of the independence of γ_1 and γ_2,

$$G_\gamma(t|\mathfrak{V}_1 + \mathfrak{V}_2) \equiv G_\gamma(t|\mathfrak{V}_1)\,G_\gamma(t|\mathfrak{V}_2). \tag{10}$$

Equation (10), valid for all $\mathfrak{V}_1$ and $\mathfrak{V}_2$, is, then, a consequence of the assumption that the distribution of cluster centers is quasi-uniform. It is easy to see that the converse is also true and that every probability generating function depending on a parameter $\mathfrak{V}$, which satisfies condition (10), will determine a quasi-uniform distribution of cluster centers.

Distributions satisfying condition (10) were studied by Paul Lévy,[10] who termed them "infinitely divisible." It follows from Lévy's work that, for every $\mathfrak{V}$,

$$G_\gamma(t|\mathfrak{V}) = [G_\gamma(t|1)]^{\mathfrak{V}}, \tag{11}$$

so that, in order to know the probability generating function of $\gamma(R)$ corresponding to an arbitrary region of volume $\mathfrak{V}$, it is sufficient to know this distribution for $\mathfrak{V} = 1$.

It can be shown[11] that the function $G_\gamma(t|1)$ must be of the form

$$G_\gamma(t|1) = \exp\left\{-h_0 + \sum_{k=1}^{\infty} h_k t^k\right\} = \exp\{h(t)\}, \text{ say}, \tag{12}$$

where $h_0, h_1, \ldots, h_k, \ldots$ are all nonnegative numbers subject to the restriction that the series $\sum_{k=1}^{\infty} h_k$ is convergent and that $\sum_{k=1}^{\infty} h_k = h_0$. Conversely, if the series formed by the nonnegative numbers $h_1, h_2, \ldots, h_k, \ldots$ is convergent and if its sum is equal to h_0, then the right-hand side of equation (12) represents a probability generating function such that, whatever $\mathfrak{V} \geqq 0$, the expression (11) is also a probability generating function. It will then satisfy equation (10). Incidentally, formulae (11) and (12) imply that, whatever $\mathfrak{V}$, the probability that the number of clusters $\gamma(R) = 0$ is necessarily a positive number.

If $h_2 = h_3 = \ldots = 0$, then equation (11) reduces to

$$G_\gamma(t|\mathfrak{V}) = e^{-\mathfrak{V}h_0(1-t)}, \tag{13}$$

which is the probability generating function of the Poisson distribution. On the other hand, if at least one of the h_k, for $k \geqq 2$, is positive, then, whatever $\mathfrak{V}$, the number of

[10] *Théorie de l'addition des variables aléatoires* (Paris: Gauthier-Villars, 1937).

[11] J. Neyman and E. L. Scott, in preparation.

cluster centers $\gamma(R)$ does not follow the Poisson law. However, it is interesting to notice that the most general distribution of $\gamma(R)$ is, in a sense, reducible to Poisson distributions.

To see this, denote by $\mu_1, \mu_2, \ldots, \mu_k, \ldots$ a sequence of completely independent Poisson variables with expectation of μ_k equal to $\mathfrak{B}h_k$ for $k = 1, 2, 3, \ldots$. Finally, let

$$\mu_0 = \sum_{k=1}^{\infty} k\mu_k . \tag{14}$$

Because of the independence of the μ_k, the probability generating function of μ_0 is equal to the product of the probability generating functions of products $k\mu_k$. Since the probability generating function of μ_k has the form (13), with h_k replacing h_0, it is obvious that the probability generating function of the product $k\mu_k$ is given by

$$e^{-\mathfrak{B}h_k(1-t^k)} . \tag{15}$$

It follows that the probability generating function of μ_0 is

$$\exp\left[-\mathfrak{B}\sum_{k=1}^{\infty} h_k(1-t^k)\right], \tag{16}$$

which, because of formulae (11) and (12) and because $h_0 = \sum_{k=1}^{\infty} h_k$, coincides with $G_\gamma(t|\mathfrak{B})$. As a result of all this, we may state that the most general quasi-uniform distribution of cluster centers requires that the number $\gamma(R)$ be distributed as is the sum (15) of the Poisson variable μ_1 plus double the Poisson variable μ_2 plus triple the Poisson variable μ_3 and so forth.

Formulae (11) and (12) imply that, if the expectation of $\gamma(R)$ exists, then the function $h(t)$ is differentiable at $t = 1$ and $E[\gamma(R)] = \mathfrak{B}h'(1)$, so that the derivative $h'(1)$ represents the average number of cluster centers per unit volume.

As mentioned above, when it is given that an arbitrary region R of volume $\mathfrak{B} > 0$ contains exactly $n > 0$ cluster centers, postulates 1, 2, and 3 imply that the co-ordinates U_i, V_i, and W_i of the ith cluster center are random variables. The joint probability distribution of all such triplets has interesting properties, some of which depend on the function $h(t)$. One of the properties is needed in the present paper and will be quoted without proof as follows.

Whatever be the quasi-uniform distribution of cluster centers and whatever be the number n $\geqq 1$ *of cluster centers known to be contained in a region* R *of finite volume* $\mathfrak{B} > 0$, *the conditional probability density function, given* n, *of the co-ordinates* U, V, W *of any one of these cluster centers taken separately is constant over the region* R *and, therefore, is equal to* $1/\mathfrak{B}$.

Another property of the distribution which is of particular interest may be stated as follows.

Given that the region R *of volume* $\mathfrak{B} > 0$ *contains exactly* n > 1 *cluster centers, say* C_1, $C_2, \ldots, C_n$, *such that* $P\{\gamma(R) = n|\mathfrak{B}\} > 0$, *the conditional probability* π_k *that exactly* k *out of the cluster centers* $C_2, C_3, \ldots, C_n$ *will coincide with* C_1 *is given by the formula,*

$$\pi_k = \frac{(k+1)\, h_{k+1} P\{\gamma(R) = n-k-1 \,|\, \mathfrak{B}\}}{\sum_{m=0}^{n-0} (m+1)\, h_{m+1} P\{\gamma(R) = n-m-1 \,|\, \mathfrak{B}\}}, \tag{17}$$

$$\text{for} \quad k = 0, 1, \ldots, n-1 .$$

It is seen that, if any of the numbers $h_2, h_3, \ldots, h_n$ is positive, then this probability may be positive. In particular, if $h_1 = h_2 = \ldots = h_{n-1} = 0$, but $h_n > 0$, then $\pi_{n-1} = 1$. This means that, in this case, all the n cluster centers present in R will coincide or, in other words, that, instead of n distinct cluster centers, there will be just one multiple cluster center of multiplicity n.

The final picture, then, of the quasi-uniform distribution of cluster centers is that, in the general case, it contains a Poisson distribution of *single* clusters with the expected number of cluster centers per unit volume equal to h_1. In addition, it contains a Poisson distribution, independent of the first, of *double* clusters with the expectation of the number of centers per unit volume equal to h_2. In addition, there will be Poisson distributions of cluster centers of *triple*, *quadruple*, etc., clusters, all completely independent. If $h_k = 0$, then there will be no clusters of multiplicity k. In particular, if $h_2 = h_3 = \ldots = 0$, then all the clusters will be single clusters.

Proofs of these properties of the general quasi-uniform distribution must be relegated to a separate publication in a statistical journal.[11] Some of these properties are foreshadowed in the work of Paul Lévy already quoted.

V. JOINT DISTRIBUTION OF THE NUMBERS N_1 AND N_2 OF GALAXIES VISIBLE WITHIN TWO ARBITRARY SOLID ANGLES

Consider two arbitrary regions, ω_1 and ω_2, overlapping or not, which will be photographed on idealized plates. The same letters ω_1 and ω_2 will be used to denote the two solid angles, with vertices at the observer, corresponding to these regions. The observational setups for the two regions may be the same or not, so that the limiting magnitude m_1 of the photograph of ω_1 need not be the same as the limiting magnitude m_2 of the photograph of ω_2. For convenience, we shall assume $m_1 \leqq m_2$.

Denote by N_1 and N_2 two random variables representing, respectively, the numbers of galaxies visible on photographs taken over ω_1 and ω_2. The purpose of this section is to deduce the joint probability generating function $G_{N_1, N_2}(t_1, t_2)$ of N_1 and N_2 as implied by the eight structural postulates enumerated above.

Briefly, the method used consists in dividing the whole space into an infinity of bounded regions of convenient shape. Each region will contain a certain number of cluster centers. Galaxies from the corresponding clusters may fall and be visible either in ω_1 or in ω_2 or in both. Thus N_1 and N_2 are represented as sums of contributions from each of the regions contemplated. Further, the contribution of each particular region is split into as many components as there are cluster centers in this region, so that N_1 and N_2 appear as double sums. The advantage of this procedure is that, according to the postulates adopted, the contribution to N_1 of any one particular cluster is independent of that of any other cluster, so that N_1 appears as the sum of completely independent components. The same applies to N_2. It is noteworthy, however, that the contributions to N_1 and N_2 from any one cluster are dependent.

After this general description, we may proceed to details. Let Δ be an arbitrary positive number which we shall later make tend to zero. Divide the whole space into an infinity of equal cubes $R_1, R_2, \ldots, R_j, \ldots$, of dimensions Δ, by passing three sequences of planes distant by Δ, each sequence parallel to one of the co-ordinate planes. The cubes R_j will be described as the elementary cubes of dimensions Δ. The order in which they are numbered is immaterial. Obviously, the volume is the same, equal to Δ^3.

Simplifying the notation adopted earlier, we shall use the symbol γ_j to denote the number of cluster centers in R_j. Assume for a moment that $\gamma_j > 0$ and consider the kth cluster center in R_j. Denote by N_{1jk} the number of galaxies visible in ω_1 which belong to the kth cluster centered in R_j. Similarly, N_{2jk} will denote the number of galaxies visible in ω_2 which belong to the same kth cluster centered in R_j. In further work it will be convenient to consider the variables N_{1jk} and N_{2jk} for $k = 0$. In order to do so, we shall adopt the convention $N_{1j0} = N_{2j0} \equiv 0$.

With this convention, N_1 and N_2 may be represented as sums of infinitely many components of the type, say,

$$N_{ij} = \sum_{k=0}^{\gamma_j} N_{ijk} \quad \text{for} \quad i = 1, 2 . \tag{18}$$

Obviously, N_{ij} represents the contribution to N_i of all the clusters centered in R_j and is equal to the sum on the right of formula (18), whether γ_j is zero or not. The postulates adopted imply that the couples of random variables (N_{11}, N_{21}), (N_{12}, N_{22}), . . . , (N_{1j}, N_{2j}), . . . are completely independent. Taking this into account and noting that

$$N_i = \sum_{j=1}^{\infty} N_{ij} , \quad \text{for} \quad i = 1, 2 , \tag{19}$$

we may write

$$G_{N_1, N_2}(t_1, t_2) = \prod_{j=1}^{\infty} G_{N_{1j}, N_{2j}}(t_1, t_2) . \tag{20}$$

Thus, in order to compute the joint probability generating function of N_1 and N_2, it is sufficient to compute that of the jth components (18) and to evaluate the infinite product on the right-hand side of equation (20).

In order to compute the probability generating function of N_{1j} and N_{2j}, we start with the definition and then apply formula (2), which connects absolute and conditional expectations:

$$G_{N_{1j}, N_{2j}}(t_1, t_2) = E\,(t_1^{N_{1j}} t_2^{N_{2j}}) = E\,[E\,(t_1^{N_{1j}} t_2^{N_{2j}} \mid \gamma_j)\,] . \tag{21}$$

If γ_j is zero, then $N_{1j} = N_{2j} = 0$. Otherwise, N_{1j} and N_{2j} are sums of the same number γ_j of components N_{1jk} and N_{2jk}, respectively. Each pair (N_{1jk}, N_{2jk}) represents the contributions to N_1 and N_2 from the kth cluster centered in R_j. The postulates adopted imply that all the γ_j pairs are completely independent. Also, the distribution of each of the γ_j clusters is exactly the same. Therefore, the joint distribution of each pair (N_{1jk}, N_{2jk}) must be the same, coinciding with that of (N_{1j1}, N_{2j1}), say. For these reasons, whether γ_j has a positive value or is zero,

$$E\,(t_1^{N_{1j}} t_2^{N_{2j}} \mid \gamma_j) = [G_{N_{1j1}, N_{2j1}}(t_1, t_2)\,]^{\gamma_j} ; \tag{22}$$

and formula (21) may be rewritten as

$$G_{N_{1j}, N_{2j}}(t_1, t_2) = E\{\,[G_{N_{1j1}, N_{2j1}}(t_1, t_2)\,]^{\gamma_j}\} . \tag{23}$$

Here $G_{N_{1j1}, N_{2j1}}(t_1, t_2)$ represents the joint probability generating function of the contributions N_{1j1} to N_1 and N_{2j1} to N_2 of any one of the clusters known to have its center in the elementary cube R_j. Referring to the definition of the probability generating function of the variable γ_j corresponding to R_j, it is easy to see that formula (23) implies

$$G_{N_{1j}, N_{2j}}(t_1, t_2) = G_\gamma\,[G_{N_{1j1}, N_{2j1}}(t_1, t_2) \mid \Delta^3] \tag{24}$$

or, using formulae (11) and (12),

$$G_{N_{1j}, N_{2j}}(t_1, t_2) = \exp\{\,\Delta^3 h\,[G_{N_{1j1}, N_{2j1}}(t_1, t_2)\,]\,\} . \tag{25}$$

Thus, in order to compute the joint probability generating function of N_{1j} and N_{2j}, it is sufficient to compute that of a single pair of components (N_{1j1}, N_{2j1}) and substitute the result for the argument of the probability generating function of γ_j.

Using formulas (20) and (25), we have

$$G_{N_1, N_2}(t_1, t_2) = \exp\{\Delta^3 \sum_{j=1}^{\infty} h[G_{N_{1j1}, N_{2j1}}(t_1, t_2)]\}. \tag{26}$$

The general form of $G_{N_{1j1}, N_{2j1}}(t_1, t_2)$ will be obtained in the next section. Here we shall notice that the series on the right-hand side ot formula (26) need not be convergent. However, the cases of divergence are easily interpreted. Owing to the particular meaning (3) of a probability generating function and because $h(t) \leqq 0$, if the series diverges at $t_1 = t_2 = 0$, this means simply that $P\{(N_1 = 0)(N_2 = 0)\} = 0$. If the same series diverges for $t_1 t_2 > 0$, then all the probabilities $P\{(N_1 = n_1)(N_2 = n_2)\} = 0$ for *all* values of $n_1, n_2 = 0, 1, 2, \ldots$. This is interpreted to mean that, in this particular case, the probability that there will be an infinite number of visible galaxies in at least one of the solid angles ω_1 and ω_2 is unity. While these and similar cases are theoretically possible, since they do not correspond to empirical facts, we shall consider only such functions $h(t)$, $G\gamma(t), f(\eta)$, and $\theta(\xi)$ with which the series on the right is convergent for all values of t_1 and t_2, $0 \leqq t_1, t_2 \leqq 1$, and, therefore, formula (26) is different from zero.

VI. FINAL FORM OF $G_{N_1, N_2}(t_1, t_2)$

The probability generating function of the variables N_{1j1}, N_{2j1} is obtained by using certain considerations of a character somewhat different from those above. For this reason it was thought useful to treat the problem in a separate section. The elementary cube R_j may contain a number of cluster centers. However, we are now interested in only one of them, namely, the first. The letters U, V, W will be used to denote the co-ordinates of its center so long as these co-ordinates remain random. The letters u, v, w will be used to denote the particular values that U, V, and W may assume. The random variable representing the number of galaxies in the cluster will be denoted by ν. We begin with the definition of the probability generating function and apply formula (2) twice. First we write

$$G_{N_{1j1}, N_{2j1}}(t_1, t_2) = E(t_1^{N_{1j1}} t_2^{N_{2j1}}) = E[E(t_1^{N_{1j1}} t_2^{N_{2j1}} | U, V, W)]. \tag{27}$$

Next we consider a fixed system of values u, v, w and consider

$$E(t_1^{N_{1j1}} t_2^{N_{2j1}} | u, v, w) = E[E(t_1^{N_{1j1}} t_2^{N_{2j1}} | u, v, w, \nu)]. \tag{28}$$

In order to proceed further, we consider the mth galaxy of the cluster and two random variables α_m and β_m defined as follows: If the galaxy falls in the solid angle ω_1 and is visible in it, then $\alpha_m = 1$. Otherwise, $\alpha_m = 0$. If the mth galaxy falls in the solid angle ω_2 and is visible there, then $\beta_m = 1$. Otherwise, $\beta_m = 0$. Obviously, the random variables α_m and β_m are dependent. Because of postulate 6, the joint distribution of (α_m, β_m) is the same for all $m = 1, 2, \ldots, \nu$. Also, because of postulate 7, all the couples (α_1, β_1), $(\alpha_2, \beta_2), \ldots, (\alpha_\nu, \beta_\nu)$ are completely independent. Finally, given a fixed value of $\nu > 0$,

$$N_{1j1} = \sum_{m=1}^{\nu} \alpha_m, \qquad N_{2j1} = \sum_{m=1}^{\nu} \beta_m. \tag{29}$$

Repeating the reasoning which now must be familiar to the reader, we find that

$$E(t_1^{N_{1j1}} t_2^{N_{2j1}} | U, V, W) = G_\nu[G_{\alpha_1, \beta_1}(t_1, t_2 | U, V, W)]. \tag{30}$$

In postulate 8 we introduced the probability that the apparent magnitude of a galaxy at distance ξ from the observer will be less than the limiting magnitude of the idealized plate. Because the observational setups used to photograph the regions ω_1 and ω_2 need not be the same, it is necessary to make a distinction between $\theta_1(\xi)$ and $\theta_2(\xi)$ corresponding to the setups to be used for regions ω_1 and ω_2, respectively.

In order to compute the probability generating function of α_1 and β_1, given that the co-ordinates of the corresponding cluster are u, v, w, denote by $\omega_1\omega_2$ the common part of the solid angles ω_1 and ω_2 and introduce the following probabilities.

Let $p_1(u, v, w)$ denote the probability that a galaxy from a cluster centered at (u, v, w) will be visible on the plate taken over the region ω_1 but not on the plate taken over ω_2. Because we have assumed that the limiting magnitude $m_1 \leqq m_2$, we have

$$p_1(u, v, w) = \iiint_{\omega_1 - \omega_1\omega_2} f(\eta)\, \theta_1(\xi)\, dx\, dy\, dz\,. \tag{31}$$

Let $p_2(u, v, w)$ denote the probability that a galaxy from a cluster centered at (u, v, w) will be visible on the plate taken over ω_2 but not on that taken over ω_1. Obviously,

$$p_2(u, v, w) = \iiint_{\omega_2 - \omega_1\omega_2} f(\eta)\, \theta_2(\xi)\, dx\, dy\, dz + \iiint_{\omega_1\omega_2} f(\eta)\, [\theta_2(\xi) - \theta_1(\xi)]\, dx\, dy\, dz\,. \tag{32}$$

Finally, let $p_3(u, v, w)$ denote the probability that a galaxy from a cluster centered at (u, v, w) will be visible on both plates,

$$p_3(u, v, w) = \iiint_{\omega_1\omega_2} f(\eta)\, \theta_1(\xi)\, dx\, dy\, dz\,. \tag{33}$$

As mentioned, the present general statement of the problem is meant to apply to two kinds of practical situations. In the first we contemplate counts of nebulae in nonoverlapping regions photographed with the same instrument, so that $\theta_1(\xi) \equiv \theta_2(\xi)$. In this case $\omega_1\omega_2$ is empty, and the integrals over this region are zero. The second situation contemplated is that in which the photographs are taken with different instruments. The most important case of this category is when ω_1 and ω_2 coincide and thus coincide with their common part $\omega_1\omega_2$. In this case the integrals taken over $\omega_1 - \omega_1\omega_2$ and over $\omega_2 - \omega_1\omega_2$ are equal to zero, so that, in particular, $p_1(u, v, w) = 0$.

Using the probabilities (31), (32), and (33), the probability generating function of α_1 and β_1 is obtained easily. In so doing, for the sake of brevity, we shall omit the reference to (u, v, w) from the symbols of the three probabilities. We obtain

$$G_{\alpha_1, \beta_1}(t_1, t_2) = 1 - p_1(1 - t_1) - p_2(1 - t_2) - p_3(1 - t_1 t_2) \tag{34}$$

and, therefore, using formula (30),

$$E(t_1^{N_{1j1}} t_2^{N_{2j1}} \mid u, v, w) = G_\nu [1 - p_1(1 - t_1) - p_2(1 - t_2) - p_3(1 - t_1 t_2)]\,. \tag{35}$$

In order to obtain the probability generating function of N_{1j1} and N_{2j1}, we use formula (27) and the result quoted at the end of Section IV to the effect that, given that a cluster center is contained in the region R_j of volume $\Delta^3 > 0$, the probability density of the co-ordinates of the cluster center is constant over R_j and equal to $1/\Delta^3$. Thus we have

$$G_{N_{1j1}, N_{2j1}}(t_1, t_2) = \frac{1}{\Delta^3} \iiint_{R_j} G_\nu [1 - p_1(1 - t_1) - p_2(1 - t_2) - p_3(1 - t_1 t_2)]\, du\, dv\, dw\,. \tag{36}$$

This formula may now be substituted in formula (26) to obtain the probability generating function of N_1 and N_2. However, before doing so, it is convenient to use the assumption that the function $f(\eta)$, representing the probability density of the co-ordinates of a galaxy, and also $\theta_1(\xi)$ and $\theta_2(\xi)$ are continuous and to apply the mean-value theorem to formula (36). Namely, the region R_j must contain at least one point, say (u_j, v_j, w_j) such that the value of formula (36) is equal to the integrand evaluated at this particular point. Denote

$$p_i(u_j, v_j, w_j) = p_{ij} \quad \text{for} \quad i = 1, 2, 3 . \tag{37}$$

Then

$$G_{N_{1j1}, N_{2j1}}(t_1, t_2) = G_\nu[1 - p_{1j}(1-t_1) - p_{2j}(1-t_2) - p_{3j}(1-t_1t_2)] . \tag{38}$$

Substituting this result in equation (25), we obtain

$$G_{N_1, N_2}(t_1, t_2) = \exp\left\{\Delta^3 \sum_{j=1}^{\infty} h(G_\nu[1 - p_{1j}(1-t_1) - p_{2j}(1-t_2) - p_{3j}(1-t_1t_2)])\right\} . \tag{39}$$

This formula is valid, irrespective of the value of Δ. Using the assumption that the series on the right is convergent for all values of t_1 and t_2, it is easy to see that, as $\Delta \to 0$, the series in the right-hand side, multiplied by Δ^3, converges to a triple integral over the whole space, so that

$$G_{N_1, N_2}(t_1, t_2) = e^{\psi(t_1, t_2)} \tag{40}$$

with

$$\psi(t_1, t_2) = \iiint_{-\infty}^{+\infty} h\{G_\nu[1 - p_1(1-t_1) - p_2(1-t_2) - p_3(1-t_1t_2)]\}\, du\, dv\, dw , \tag{41}$$

where the symbols p_1, p_2, p_3 have the meaning defined in formulae (31), (32), and (33), respectively.

Formulae (40) and (41) determine the probability generating function of N_1 and N_2 in its final form. With the functions $h(t)$, $G_\nu(t)$, $f(\eta)$, and $\theta(\xi)$ not specialized, formulae (40) and (41) reflect only the general structure of the universe contemplated and not the details considered unimportant. For this reason, the properties of formula (40) which are independent of the unspecified functions are of particular interest.

VII. MOMENTS OF N_1 AND N_2

In the present section we use formula (41) to deduce general expressions for the moments and product moments of N_1 and N_2 in terms of the functions $h(t)$, $G_\nu(t)$, $f(\eta)$, and $\theta(\xi)$. The formulae so deduced reveal certain interesting properties of the variables N_1 and N_2 which are independent of the particular forms that the four functions may possess. Of course, in order to deal with these expressions, it is necessary to assume that the moments exist. This assumption appears very plausible and will be adopted from now on. More specifically, we shall limit our consideration to functions $h(t)$ and $G_\nu(t)$ which are indefinitely differentiable at $t = 1$.

In Section II we mentioned that the derivatives of the probability generating function evaluated at $t_1 = t_2 = \ldots = t_s = 1$ are connected with the moments of the variables concerned. A similar statement applies to the natural logarithm of the probability generating function. Since formula (41) involves two such logarithms, namely, $\psi(t_1, t_2)$ and $h(t)$, further work will be simplified if the connection between the moments and the derivatives of the logarithm of the probability generating function is given explicitly. The derivation of the formulae is elementary and therefore not given here. The formulae reproduced refer to $\psi(t_1, t_2)$, but the reader will have no difficulty in applying them to $h(t)$ by analogy. Although not explicitly marked, all the derivatives appearing in the formulae are assumed to be evaluated at both arguments equal to unity. The symbol σ_{rs} will be used to denote the central product moment of order r with respect to N_1 and of order s with respect to N_2, so that

$$\sigma_{rs}=E\{[N_1-E(N_1)]^r[N_2-E(N_2)]^s\} \tag{42}$$

for $r, s = 0, 1, 2, \ldots$. With this notation, σ_{20} stands for the variance of N_1, etc. We have

$$\frac{\partial\psi}{\partial t_i}=E(N_i), \qquad i=1\ 2, \tag{43}$$

$$\frac{\partial^2\psi}{\partial t_1^2}=\sigma_{20}-E(N_1), \tag{44}$$

and a similar expression for $\partial^2\psi/\partial t_2^2$:

$$\frac{\partial^2\psi}{\partial t_1\partial t_2}=\sigma_{11}, \tag{45}$$

$$\frac{\partial^3\psi}{\partial t_1^3}=\sigma_{30}-3\sigma_{20}+2E(N_1), \tag{46}$$

$$\frac{\partial^3\psi}{\partial t_1^2\partial t_2}=\sigma_{21}-\sigma_{11}, \text{ etc.} \tag{47}$$

The product moments σ_{rs} may be estimated from empirical data.

We now return to the expression for $\psi(t_1, t_2)$ given in formula (41) and compute the successive partial derivatives with respect to t_1 and t_2, substituting each time $t_1 = t_2 = 1$. In so doing, it will be convenient to write h_j and G_j for the jth derivatives of $h(t)$ and $G_\nu(t)$, respectively, evaluated at $t = 1$, and to write R_{kmn} for the triple integral

$$R_{kmn}=\iiint_{-\infty}^{+\infty}(p_1+p_3)^k(p_2+p_3)^m p_3^n\,du\,dv\,dw. \tag{48}$$

Notice that the substitution of $t_1 = t_2 = 1$ into formula (41) reduces the arguments of both G_ν and h to unity.

Taking derivatives of the first order, we obtain

$$E(N_1)=A_1R_{100}, \qquad E(N_2)=A_1R_{010}, \tag{49}$$

where

$$A_1=h_1G_1. \tag{50}$$

Derivatives of the second order yield the following three formulae:

$$\sigma_{20}-E(N_1)=A_2R_{200}, \tag{51}$$

$$\sigma_{11}=A_2R_{110}+A_1R_{001} \tag{52}$$

$$\sigma_{02}-E(N_2)=A_2R_{020}, \tag{53}$$

where, for the sake of brevity,

$$A_2=h_2G_1^2+h_1G_2. \tag{54}$$

Taking derivatives of the third order, we obtain

$$\sigma_{30} - 3\sigma_{20} + 2E(N_1) = A_3 R_{300}, \tag{55}$$

$$\sigma_{21} - \sigma_{11} = A_3 R_{210} + A_2 R_{101}, \tag{56}$$

$$\sigma_{12} - \sigma_{11} = A_3 R_{120} + A_2 R_{011} \tag{57}$$

$$\sigma_{03} - 3\sigma_{02} + 2E(N_2) = A_3 R_{030}, \tag{58}$$

where

$$A_3 = h_3 G_1^3 + 3 h_2 G_1 G_2 + h_1 G_3 . \tag{59}$$

The sequence of these equations can be continued indefinitely. However, the foregoing formulae are sufficient to give an idea of the nature of all the equations obtainable in this manner. The expressions on the left-hand sides are combinations of moments of N_1 and N_2 and, therefore, can be estimated from empirical data. The expressions on the right-hand side are products or sums of products of quantities of two different kinds. First, there are quantities A_j expressible in terms of h_i and G_i and thus, ultimately, in terms of moments of the variables γ and ν. It is important to notice that the A_j's are independent of the internal structure of the clusters, which is governed by the function $f(\eta)$, and of the distribution of apparent magnitude of galaxies, as reflected in the function $\theta(\xi)$. Also, the quantities A_j are independent of the regions ω_1 and ω_2.

The other kind of quantity appearing on the right-hand side of the equations are the triple integrals R_{kmn}. The simplest of these integrals, namely, R_{100}, R_{010}, and R_{001}, depend on the function $\theta(\xi)$ and on the regions ω_1 and ω_2 but not on anything else. In fact,

$$R_{100} = \iiint_{\omega_1} \theta_1(\xi)\, dx\, dy\, dz, \tag{60}$$

$$R_{010} = \iiint_{\omega_2} \theta_2(\xi)\, dx\, dy\, dz, \tag{61}$$

$$R_{001} = \iiint_{\omega_1 \omega_2} \theta_1(\xi)\, dx\, dy\, dz. \tag{62}$$

It will be sufficient to prove only one of these formulae, for example, the first. Referring to formulae (31) and (33), we have

$$p_1 + p_3 = \iiint_{\omega_1} f(\eta)\, \theta_1(\xi)\, dx\, dy\, dz, \tag{63}$$

and R_{100} is defined as the integral of this quantity with respect to u, v, w taken from $-\infty$ to $+\infty$ for each of the three variables. Now the variables u, v, w enter into formula (63) only through η. Therefore, changing the order of integration, we obtain

$$R_{100} = \iiint_{\omega_1} \left\{ \theta_1(\xi) \iiint_{-\infty}^{+\infty} f(\eta)\, du\, dv\, dw \right\} dx\, dy\, dz; \tag{64}$$

and it is easy to show that the internal triple integral is equal to unity regardless of the particular form of the function

$$f(\eta) = f\{[(x-u)^2 + (y-v)^2 + (z-w)^2]^{1/2}\}, \tag{65}$$

which, it will be remembered, represents the probability density of the co-ordinates of a galaxy belonging to a cluster centered at (u, v, w). This particular property of the function $f(\eta)$ implies that, whatever be u, v, w, the integral of $f(\eta)$ with respect to x, y, z, taken from $-\infty$ to $+\infty$ for each variable, must be equal to unity. Making the substitution

$$u = x + x' , \qquad y = y + y' , \qquad w = z + z' , \tag{66}$$

where x', y', and z' are new variables of integration, the internal integral in formula (64) is reduced to

$$\iiint_{-\infty}^{+\infty} f\{ [x'^2 + y'^2 + z'^2]^{1/2} \} dx' dy' dz' = 1 , \tag{67}$$

which, combined with formula (64), proves formula (60). Formulae (61) and (62) are proved in exactly the same manner.

If it is assumed that the apparent luminosity of the galaxies varies inversely as the square of their distance (thus, if one ignores "red shift") and that the absolute magnitude of the galaxies is a random variable independent of the distance, then the functions $\theta_1(\xi)$ and $\theta_2(\xi)$ are relatively simple, and the integrals (59)–(61) can be evaluated as explicit functions of the limiting apparent magnitude. The resulting formula for $E(N_i)$ is well known.[12] Ordinarily, it is deduced on the assumption that the luminous objects (stars or galaxies) are distributed in space Poisson-wise. It is interesting that the same formula appears to hold on the more general assumption that the galaxies appear in clusters and that the cluster centers are quasi-uniformly distributed in space.

If the effect of red shift is taken into account, then the function $\theta(\xi)$ becomes complicated, and the values of R_{100}, R_{010}, and R_{001} must be obtained by numerical integration. The integrals R_{kmn} with $k + m + n > 1$ depend essentially not only on $\theta(\xi)$ but also on $f(\eta)$. In general, their evaluation appears to require numerical integration.

Perusing formulae (49)–(58), we see that they can be combined so as to eliminate the quantities A_j. The equations resulting from such eliminations connect the integrals R_{kmn} directly with the moments of N_1 and N_2. Thus, for example, combining formulae (49)–(53), we obtain, say,

$$\begin{aligned} Q(\omega_1, \omega_2) &\equiv \frac{R_{110}}{\{R_{200} R_{020}\}^{1/2}} \\ &= \frac{\sigma_{11} R_{100} - E(N_1) R_{001}}{R_{100}\{[\sigma_{20} - E(N_1)][\sigma_{02} - E(N_2)]\}^{1/2}} , \end{aligned} \tag{68}$$

which is an equation containing only integrals R_{kmn} and the moments of N_1 and N_2. Similar equations can be obtained by eliminating A_2 and A_3 from formulae (51) and (55)–(58).

The existence of relations between the integrals R_{kmn} and the moments of N_1 and N_2 is a very important fact because it creates the possibility of studying the internal structure of the clusters, more specifically, the functions $f(\eta)$ and $\theta(\xi)$, in a manner which is independent of how the cluster centers are distributed in space and independent of the variation of the number of galaxies from one cluster to another. While details of this study must be relegated to a later paper, some indications as to how this can be achieved are given in Section IX.

Returning to the quantities A_j, we see that they represent derivatives at $t = 1$ of the function $h[G_v(t)]$. This function can be interpreted as the logarithm of the probability generating function of the random variable ζ defined to be the total number of galaxies

[12] W. M. Smart, *Stellar Dynamics* (London: Cambridge University Press, 1938), p. 263.

belonging to all the clusters which have their centers in a specified region of unit volume. Once reasonable approximations to the functions $f(\eta)$ and $\theta(\xi)$ have been determined and given a substantial set of empirical material, any number of moments A_j can be obtained from equations such as (49)–(58). The estimates of these moments can then be used to approximate the distribution of ζ. Although of interest, ζ is not what one might consider a basic concept in the structure of the universe contemplated. Instead, our interest is directed toward the variables γ and ν.

Since the moments of A_j are simple combinations of quantities h_i and G_i, which, in turn, are simple functions of the moments of γ and ν, one might hope that knowledge of the moments A_j for $j = 1, 2, \ldots, s$ would make possible the evaluation of at least a few of the moments of γ and ν taken separately. Unfortunately, this is not the case because each A_j depends on both of the quantities h_j and G_j, which do not appear in $A_1, A_2, \ldots, A_{j-1}$. Thus, even with the knowledge of the functions $f(\eta)$ and $\theta(\xi)$, unless independent information regarding the distribution of γ and ν is obtained from other sources, no separate moments of γ and ν can be determined from any finite set of moments of N_1 and N_2. The next section gives an even stronger result to the same general effect.

VIII. NONIDENTIFIABILITY OF THE DISTRIBUTION OF CLUSTER CENTERS

In the present section we shall show that the same joint distribution of N_1 and N_2 can result from an infinity of different pairs of distributions of γ and ν and that, therefore, knowledge of the joint distribution of N_1 and N_2 is not sufficient to determine the distributions of γ and ν separately.

For this purpose we return to formula (41) and notice that the integrand depends not on the functions $h(t)$ and $G_\nu(t)$ taken individually but rather on the result of substituting $G_\nu(t)$ instead of the argument in $h(t)$. Thus the probability generating function $G_{N_1, N_2}(t_1, t_2)$ depends on the nature of the function, say,

$$H(t) = h[G_\nu(t)]. \tag{69}$$

Hence, if the two functions $h(t)$ and $G_\nu(t)$ are replaced by any other two functions, say, $h^*(t)$ and $G_\nu^*(t)$, such that, however,

$$h^*[G_\nu^*(t)] \equiv h[G_\nu(t)], \tag{70}$$

then this change will not produce any change in the joint distribution of N_1 and N_2. The existence of functions $h^*(t)$ and $G_\nu^*(t)$ having the above property is assured by the following theorem.

THEOREM.—*Whatever be the (infinitely divisible) distribution of γ with probability generating function*

$$G_\gamma(t \mid \mathfrak{B} = 1) = e^{h(t)}, \tag{71}$$

where

$$h(t) = -h_0 + \sum_{k=1}^{\infty} h_k t^k, \tag{72}$$

with $h_k \geqq 0$, $k = 0, 1, \ldots$ *and* $h(1) = 0$, *and whatever be the distribution of ν, with probability generating function* $G_\nu(t)$, *there exists an infinity of Poisson laws with probability generating functions, say,*

$$G_\gamma^*(t \mid \lambda) = e^{\lambda(t-1)}, \tag{73}$$

where λ is any positive number exceeding a certain limit λ_0, and a set of corresponding probability generating functions $G_\nu^*(t \mid \lambda)$ *such that*

$$G_\gamma[G_\nu(t) \mid \mathfrak{B} = 1] \equiv G_\gamma^*[G_\nu^*(t \mid \lambda) \mid \lambda]. \tag{74}$$

Proof.—In order to prove formula (74), we use equation (73) with the parameter λ unspecified, substitute it in (74), and take the logarithms of both sides. Using formulae (69) and (71), the identity to be proved is reduced to

$$H(t) = h[G_\nu(t)] = -h_0 + \sum_{k=1}^{\infty} h_k [G_\nu(t)]^k \equiv \lambda [G_\nu^*(t \mid \lambda) - 1] \tag{75}$$

or to

$$G_\nu^*(t \mid \lambda) \equiv 1 + \frac{1}{\lambda} H(t). \tag{76}$$

It is seen that every value of $\lambda > 0$ determines a corresponding function $G_\nu^*(t \mid \lambda)$. In order to complete the proof of the theorem, we have to show that formula (76) is a probability generating function. We shall show that this is true for an infinity of values of λ—in fact, for all values of λ which exceed a certain limit λ_0.

In order that formula (76) may be a probability generating function, it must have the following properties:

(i) $$G_\nu^*(1 \mid \lambda) = 1,$$

(ii) $$G_\nu^*(0 \mid \lambda) \geqq 0,$$

(iii) $$\left. \frac{d^m G_\nu^*(t \mid \lambda)}{dt^m} \right|_{t=0} \geqq 0 \quad \text{for} \quad m = 1, 2, \ldots.$$

In order to see that formula (i) is satisfied, we substitute $t = 1$ in formula (76) and then refer to formula (75). Since $G_\nu(1)$ is necessarily unity and since $h(1) = 0$, we obtain $H(1) = 0$ and formula (i) follows.

Proceeding to property (ii), we substitute $t = 0$ in formula (76) and again refer to (75). Since $G_\nu(0) = P\{\nu = 0\} = p_0$, say, we have

$$G_\nu^*(0 \mid \lambda) = 1 + \frac{1}{\lambda} H(0), \tag{77}$$

where

$$H(0) = -h_0 + \sum_{k=1}^{\infty} h_k p_0^k, \tag{78}$$

which is easily seen to be a finite negative number. Thus, condition (ii) will be satisfied if $\lambda \geqq -H(0) = \lambda_0$, say.

In order to prove condition (iii), we notice that

$$\left. \frac{d^m G_\nu^*(t \mid \lambda)}{dt^m} \right|_{t=0} = \frac{1}{\lambda} \left. \frac{d^m H(t)}{dt^m} \right|_{t=0} = \frac{1}{\lambda} \frac{d^m}{dt^m} \sum_{k=1}^{\infty} h_k [G_\nu(t)]^k \Bigg|_{t=0}. \tag{79}$$

Since the derivatives of $G_\nu(t)$, evaluated at $t = 0$, are all nonnegative and since the constants $h_k \geqq 0$ by hypothesis, property (iii) follows.

Obviously, formula (74) implies formula (70).

In order to illustrate the bearing of this theorem, assume for a moment that the true distribution of the number γ of cluster centers per unit volume is a particular negative binomial law with probability generating function

$$G_\gamma(t \mid \mathfrak{V}=1) = [2-t]^{-1} \tag{80}$$

and that this is combined with a specified binomial distribution of the number ν of galaxies per cluster, so that

$$G_\nu(t) = [\tfrac{1}{2}(1+t)]^{1000} . \tag{81}$$

Then

$$H(t) = h[G_\nu(t)] = -\ln\{2-[\tfrac{1}{2}(1+t)]^{1000}\}, \tag{82}$$

and

$$G_\nu^*(t) = 1-\frac{1}{\lambda}\ln\{2-[\tfrac{1}{2}(1+t)]^{1000}\}, \tag{83}$$

with

$$\lambda \geqq -H(0) = \ln\{2-(\tfrac{1}{2})^{1000}\} = \lambda_0 . \tag{84}$$

It follows that the joint distribution of N_1 and N_2 resulting from the true distributions (80) and (81) will be identical with that which would be observed if the number of cluster centers per unit volume followed a Poisson distribution with expectation λ satisfying (84) combined with the distribution of the number of galaxies per cluster as determined by formula (83). Naturally, therefore, observing the joint distribution of N_1 and N_2 and nothing else, it is impossible to decide whether the number of cluster centers per unit of volume follows a Poisson law or any other quasi-uniform distribution. This, of course, does not exclude the possibility that we can decide this question on other grounds.

IX. CONCLUDING REMARKS

All the preceding pages are given to deductions from the structural postulates regarding the spatial distribution of galaxies as enumerated in Section I and, in a more precise form, in Section III. The main result of this work is represented by formulae (40) and (41), which determine the joint distribution of the numbers N_1 and N_2 of galaxies visible on idealized photographs of two arbitrary regions ω_1 and ω_2. Confrontation of the theory thus developed with empirical facts requires specialization of at least some of the four functions involved in the model. The work done in this direction will be described in a subsequent publication.[13] However, before concluding the present paper, it seems advisable to include at least a few remarks indicating how formula (41) and its consequences (49)–(58) can be used for an empirical study of the actual distribution of galaxies.

Keeping in mind the general properties of the moments of N_1 and N_2 described in Section VII, it appears advisable to divide the empirical study into two parts: the internal structure of clusters and the distribution of cluster centers.

The theoretical results which are basic for the first part of the study are exemplified by equation (68). Because of the analogy between the definition of correlation coefficient and the expression $Q(\omega_1, \omega_2)$ on the left-hand side of equation (68), we propose to describe this quantity as the *quasi-correlation* between N_1 and N_2. At least in the early stages of the empirical study, the function $\theta_i(\xi)$ will be taken from the published work of Hubble[14] or Holmberg.[15] When this is done, the right-hand side of equation (68) contains only known quantities and the moments of N_1 and N_2, for which estimates from actual counts will be substituted. The selection of ω_1 and ω_2 and of m_1 and m_2 is at our disposal and may

[13] J. Neyman, E. L. Scott, and C. D. Shane, in preparation.

[14] *Mt. W. Contr.*, No. 548; *Ap. J.*, **84**, 158, 1936.

[15] *Medd. Lunds Obs.*, Ser. II, No. 128, 1950.

be adjusted to the material available. For example, ω_1 and ω_2 may mean two $1^\circ \times 1^\circ$ squares taken k degrees apart, with N_1 and N_2 representing the numbers of galaxies in these squares up to the limiting magnitude appropriate to the observations made at the Lick Observatory or any other. Then the regions ω_1 and ω_2 do not overlap, $\theta_1(\xi) = \theta_2(\xi)$, and formula (68) giving the quasi-correlation between N_1 and N_2 reduces to

$$Q(\omega_1, \omega_2) = \frac{R_{110}}{R_{200}} = \frac{\sigma_{11}}{\sigma_{20} - E(N_1)} = Q_k \quad \text{(say)}. \tag{85}$$

If the available set of data is substantial, then empirical values, say Q_k^*, of Q_k may be obtained for a number of values of k. If the present model approximates the actual distribution of galaxies, then, *solely by adjusting the function* $f(\eta)$, the sequence of quotients R_{110}/R_{200}, computed for $k = 1, 2, \ldots$, can be made to agree with the series of numbers Q_k^*. This is true of the quasi-correlation as defined in (68) and (85) but, perhaps a little unexpectedly, no such statement can be made with regard to the ordinary correlation coefficient between N_1 and N_2, say,

$$\rho = \frac{\sigma_{11}}{\sigma_{20}} = \frac{A_2 R_{110}}{A_2 R_{200} + E(N_1)}, \tag{86}$$

because the value of ρ depends on A_2 and hence on the distribution of cluster centers in space.

Having obtained empirical quasi-correlations Q_k^*, the search for an appropriate function $f(\eta)$ may be attempted by the trial-and-error method. We choose a plausible family of probability densities, say $f(\eta, \vartheta)$, perhaps depending on a single parameter ϑ. For selected values of ϑ a sequence of values of the quotient R_{110}/R_{200} is computed for $k = 1, 2, \ldots$. Each such sequence is compared with the empirical sequence of numbers Q_k^*. Finally, the value of the parameter, say ϑ^*, is determined which provides the best agreement between the quotients R_{110}/R_{200} and the numbers Q_k^*. The corresponding probability density $f(\eta, \vartheta^*)$ is then our first approximation to the unknown function $f(\eta)$ and is open to various tests of validity.

One such test may consist in using $f(\eta, \vartheta^*)$ to compute quasi-correlations between numbers N_1 and N_2 of galaxies in the same region $\omega_1 \equiv \omega_2$ but counted to two different limiting apparent magnitudes, $m_1 < m_2$. Empirical material with which results of this kind can be compared consists of two systems of counts covering the same region of the sky.

If the empirical numbers $Q_1^*, Q_2^*, \ldots$ cannot be approximated by any choice of the function $f(\eta)$, this would indicate that the probabilistic model of the structure of the universe is essentially wrong and that it should be modified, perhaps by including the effects of extinction by interstellar clouds, etc.

Once a plausible approximation $f(\eta, \vartheta^*)$ of $f(\eta)$ is obtained, a similar method may be used to find plausible approximations to the distributions of γ and ν. A certain number of moments A_j are estimated, as outlined in Section VII. Then approximations to $G_\gamma(t \mid \mathfrak{B} = 1)$ and $G_\nu(t)$ are sought among functions with few adjustable parameters.

At this point we should make a distinction between the following two situations. First, it is possible that some a priori considerations will suggest special forms of the distributions of γ and of ν which may be approximated by probability generating functions of known form involving several adjustable parameters. In this case the quantities h_k and G_k can be represented by known functions of the same parameters and, knowing the values of the moments A_j, the best-fitting values of the parameters can be found.

If no a priori knowledge of the distribution of γ and ν is available, then the best we can do is to refer to Section VIII above and to estimate that distribution of ν which, combined with a Poisson distribution of γ, could yield the joint distribution of N_1 and N_2 that is actually observed.

The treatment of this aspect of the problem requires setting

$$h(t) = \lambda(t-1), \qquad h_1 = \lambda, \qquad h_k = 0 \quad \text{for} \quad k > 1, \tag{87}$$

where $\lambda > 0$ is an adjustable parameter. Then, for $j = 1, 2, \ldots,$

$$A_j = G_i \lambda^j ; \tag{88}$$

and, if the distribution of ν is approximated by a known distribution with several adjustable parameters, the best-fitting values of these parameters and of λ can be obtained from equations of the type of (88).

The present paper originated from conversations with Dr. C. D. Shane, director of the Lick Observatory, to whom the authors are deeply indebted. Although signed only by the present two authors, the paper should be considered as the first part of a broader study, involving not only theoretical considerations but also the analysis of empirical data, conducted jointly with Dr. Shane.