# GALAXIES, STATISTICS AND RELATIVITY

G. C. McVITTIE
UNIVERSITY OF ILLINOIS OBSERVATORY

Any investigation of the cosmological problem is bound to entail statistical considerations of some kind because of the immense number of the units of matter that are under consideration. Each such unit is itself a galaxy comparable in size with our own galaxy and populated by perhaps two or three billion stars and there are scores, possibly hundreds, of millions of such objects within range of our largest telescopes. If cosmology is the study of the entire system of galaxies, of their spatial distribution and of their relative motions, it is clear that each individual specimen cannot be observed and discussed separately. We must content ourselves with sampling procedures and attempt to deduce from the character of samples what the mechanical behaviour of the whole system may be. Nor indeed can we pick samples at random throughout the entire system, but only from that part of it which is accessible to our instruments. It appears that, even with the 200-inch telescope, the number of galaxies does not show any signs of coming to an end, countless millions lying at the extreme limit of the plates, too faint to be usefully observed but indicating that we are dealing with a part only of some unknown larger system.

As examples of the more rudimentary applications of statistics we can consider the questions of the luminosity-function of the galaxies and of the relation between red-shift and apparent magnitude, the so-called velocity-distance relation. In principle, the first question resolves itself into the determination of absolute magnitudes and the establishment of a frequency relation of absolute magnitude against numbers of galaxies. Clearly, selectivity can easily vitiate the result: if we simply pick all the brightest specimens that we can see, a completely artificial frequency distribution will result. One way of overcoming this difficulty is to adopt Bigay's method [1] which consists in selecting those galaxies that can be resolved into stars, irrespective of whether they are conspicuous or not. From the apparent and absolute magnitudes of identifiable stars in it, the absolute magnitude of a galaxy can be deduced, provided that its apparent magnitude has been measured. This is a slow and difficult process, much more difficult than is the determination of the apparent magnitude of a stellar object. However, Bigay has measured 64 galaxies that are resolvable and he concludes that their absolute magnitudes lie very closely on a Gaussian distribution curve; though a conclusion based on so small a sample is not conclusive, it can be provisionally accepted until further work, as careful and detailed as Bigay's, gives sound reasons for modifying it.

Turning now to the red-shift and apparent magnitude relation, it appears to be true that the red-shift can be measured to a much higher degree of accuracy than can the apparent magnitude of a galaxy. But to draw any conclusions of value in cosmology from this relation, the absolute magnitudes of the objects whose red-shifts are being measured must also be known. This latter problem can be solved, after a fashion, by concentrating on the giant clusters of galaxies and by assuming that the brightest members in each cluster are comparable in luminosity with those systems such as M 31 or M 81, where resolution into stars is possible. But apart from this difficulty, there are others; for example, all the galaxies of a cluster do not yield the same value for the red-shift. If the fractional red-shift in light of wavelength $l$ is $z = dl/l$, then the values of $z$ for the eight measured galaxies of Cluster 1520 + 2754 in Corona Borealis [2] range from 0.065 to 0.076, and similar ranges are found in other clusters. At the moment, the red-shift for a cluster as a whole is obtained by the simple device of taking the arithmetic mean of the values of $z$ for individual members of the cluster, but one may well wonder what statisticians think of such a process when it is remembered that the number of measurements per cluster ranges from one to eight or nine. When data are increased in number, it may well be of advantage to have a procedure laid down for us by statisticians that will deal adequately with both the spread in red-shift values within a cluster and with the comparative uncertainty of apparent magnitudes relative to red-shifts.

Important as the red-shift law is to the theoretical cosmologist, the galaxy-count is still more important, for it can be pushed to much fainter magnitudes and thus gives much better data on the nature of the system of galaxies as a whole. By a galaxy-count is meant the determination of the total number of galaxies over the entire celestial sphere visible down to some preassigned limiting magnitude. Obviously, every square degree of the sky cannot be separately photographed and the numbers counted, particularly as, to have theoretical value, the counts must be repeated for a number of limiting magnitudes. The fixing of the limiting magnitudes is a matter of great difficulty and has not yet been satisfactorily resolved. Apart from this, however, the estimation of a mean number of galaxies per square degree of the sky is complicated by the fact that the images of galaxies are not uniformly distributed over the photographic plate [3], but occur in a manner indicating that most galaxies are members of clusters. Certainly this clustering represents, from the cosmological point of view, a small-scale effect, even the largest clusters containing no more than about a thousand members. Taking the celestial sphere as a whole, the clusters themselves seem to be distributed in a random manner, there being for example no tendency to find more of them in one direction than in another. The statistical investigations of J. Neyman and others [4], [5], are interesting to the cosmologist because they suggest that significant conclusions may be drawn from galaxy-counts made to one single limiting magnitude, the same for all plates used, but without an accurate determination of this limiting magnitude. In principle, the Neyman-Scott theory implies that all galaxies are members of clusters and that the centers of clusters are distributed at random. If $C$ is a typical cluster-center then the galaxies belonging to this cluster are randomly distributed and the probability density of the galaxies $G$ belonging to the cluster with center $C$ is $f(S)$, where $S$ is a parameter specifying the separation between $C$ and $G$. Further, let us imagine that all clusters are photographed simultaneously from a single point of observation $O$, and that the limiting magnitude of the plates is $m'$; then it will be assumed that there is a probability $\theta(m', D)$ that the apparent magnitude $m$ of the galaxy $G$ is less than $m'$. The parameter $D$ presumably depends on the apparent and absolute magnitudes, $m$ and $M$, of $G$ and also on the red-shift $z$ in the light received at $O$

from $G$. With the aid of $f$ and $\theta$, all the functions required by Neyman and Scott can be built up, provided however that the element of volume $dV$ in space can also be defined. The task of the theoretical cosmologist is to define the meanings of the parameters $S$ and $D$ and of the volume-element $dV$ when the red-shift is interpreted as a Doppler displacement, so that the galaxies are regarded as being in motion.

To do this we may use the principles of either Newtonian mechanics or those of general relativity but the interpretations will not be the same in the two cases. It may be asked at this point: why is this necessary? If the red-shift versus apparent magnitude relation can be converted into a velocity-distance relation by some plausible argument, or if we interpret, for instance, the Neyman-Scott quasi correlations for numbers of galaxies counted in neighbouring areas of the celestial sphere by means of some intuitive argument based on Euclidean geometry and Newtonian kinematics, why is anything more needed? The answer is this: The experience of the past twenty years has amply demonstrated that conclusions drawn in this way depend in a comparatively unimportant manner on the observational data, and are almost entirely conditioned by the unconscious assumptions contained in the plausible or intuitive arguments. These arbitrary factors may be slipped in by making corrections to apparent magnitudes that depend on the red-shift, by the definition of the velocity of recession as a function of the red-shift, or by the interpretation of apparent magnitude in terms of distance. The theoretician today must therefore be aware of these pitfalls and must carefully examine each step in his argument so as to avoid as best he can the introduction of postulates that can later be shown to be the real determining factors in his results.

At the outset, we must decide whether we are going to use the mechanics and gravitational theory of Newton or the theory of relativity. It is sometimes argued that a linear velocity-distance relation indicates that gravitational effects do not control the motion of the galaxies at the present time, but this is a fallacy since even Newtonian cosmology shows that a linear velocity-distance relation is compatible with gravitational interaction. Now it is known that relativistic gravitational theory is more nearly in agreement with observation than is Newtonian, particularly when high speeds are involved. For example, the rotation of the perihelion of Mercury predicted by relativity theory, and not by Newtonian, may be said to arise because the speed of the planet in its orbit (48 km/sec) is already a sufficiently large fraction of the velocity of light (300,000 km/sec) to make Newtonian theory uncertain in its application. And if the red-shift is interpreted by the classical Doppler formula, we are dealing not with 48 km/sec but with speeds up to 60,000 km/sec. This consideration alone should warn us that relativity theory is the appropriate tool to use in cosmology; if it turns out that we have been overaccurate, we can always return to the Newtonian scheme by methods of approximation. Another advantage is that relativity is still a comparatively unfamiliar theory and that we are therefore forced to think out carefully each step in our argument, thereby avoiding the fatal shortcuts which long familiarity with Newtonian mechanics and Euclidean geometry makes almost unavoidable.

If then we are going to interpret the Neyman-Scott statistical theory in terms of a system of galaxies that are in motion and are subject to their mutual gravitational interaction, it is relativity theory that must be used. Since the problem of the gravitational motion of two or more mass-points in relativity theory is still unsolved, we are forced at the outset to adopt an indirect device and replace the discrete galaxies by a continuous distribution of matter so adjusted as to have the same gravitational effect. To think that we can escape from this artifice by returning to Newtonian gravitational theory is an illusion, because the three-, or more, body problem is there as unsolved as is the two-body problem in relativity, and Newtonian cosmology also employs the continuous distribution of matter postulate. Adopting this solution of the gravitational part of our problem, the next question is the definition of the separation $S$ of the galaxy $G$ and the center $C$ to which it belongs. The answer seems obvious: $S$ is the distance between $G$ and $C$; but what do we mean by this distance? Is it distance in Euclidean space? Even if this is so, is it the distance between $G$ and $C$ at the moment when the light by which $G$ is seen at $O$ left $G$, or at the moment when the light arrived at $O$? We do not in any case know that space is Euclidean, but we do know that, even if it is, the distance between two objects is not an invariant in relativity theory. This was one of the earliest results of special relativity and was interpreted in the early days of the theory as a property of rigid measuring rods, the so-called Fitzgerald contraction. A measuring scale was supposed to shrink in length in the direction of its motion, but this rather far-fetched interpretation is unnecessary; since distance is no longer an invariant, as it was in Newtonian mechanics, it becomes a quantity whose value depends on the observational procedure by which it is to be measured and on the computations to which the measurements are subjected. Different observational procedures, and different methods of computation can, and in general do, lead to different values for the distance between two objects. Suppose then that we have defined a time-variable and certain spatial coordinates by the methods used in relativistic cosmology, into which I cannot enter here but which can be found in the textbooks. The spatial coordinate-position of the observer $O$, the galaxy $G$ and the cluster-center $C$ can then be assigned; suppose also that the light which left $G$ at time $t'$ reaches $O$ at time $T$ and that $t$ is the time at which light would have to leave $C$ in order to arrive at $O$ at the same instant $T$. Then the rules of relativity permit us to define a kind of distance that I call the space-interval, which is an invariant and depends on the spatial coordinates of $G$, $C$ and $O$ and also on the instants $t'$, $t$ and $T$, and it is this space-interval that I identify with $S$. The procedure would take too long to describe in detail but I might mention two points: the first is that the space-coordinates are so chosen that $G$, $C$ and $O$ have fixed spatial coordinates in spite of the fact that all three are in relative motion. At first sight this may seem to be a contradiction in terms but this is not so; the trick employed is to use co-moving systems in terms of which the spatial coordinate-mesh expands so as to follow the motion of all galaxies and cluster-centers relative to $O$. The device can be used whether the motion takes place at constant or at variable velocity provided that at successive instants, the relative configuration remains similar to itself or, in other words, when a certain kind of uniformity is postulated. The second point is that the times $t'$ and $t$ are expressible in terms of $T$ and of the radial coordinates of $G$ and $C$, the last two being fixed numbers in virtue of the co-moving character of the coordinate system. Thus, not only have we introduced an invariant definition of the separation of $G$ and $C$ to be used as an independent variable in the probability density function $f$, but we have in a sense reduced the description of a moving system of galaxies and cluster-centers to that of a statical system at the particular instant of observation $T$.

The times involved in cosmology are, of course, of the order of millions or of billions of years and therefore all the observations performed on the galaxies in the past thirty years can be regarded as occurring at one and the same instant $T$. A satisfactory definition of the volume element $dV$ is also possible through the reduction to a statical problem and integrations over all galaxies and cluster-centers visible at $O$ at the time $T$ can then be carried out in safety, the formulas automatically omitting from consideration those clusters which, for example, are receding so quickly that the light they emit could not reach $O$.

Turning next to the function $\theta$ defining the probability that the apparent magnitude of the galaxy $G$ shall be less than the plate limit $m'$, it would seem appropriate to define the parameter $D$ (measured in parsecs) by the formula

$$\log_{10} D = 0.2(m - M) + F(z) + 1\,, \tag{1}$$

where $M$ is the absolute magnitude of $G$ and $F(z)$ represents certain allowances that have to be made because of the red-shift. This definition implies that $D$ is the luminosity-distance of $G$, that is, a distance such that the intensity of the light emitted by $G$ falls off inversely as the square of $D$. Incidentally, one of the hardest tasks that faces the theoretical cosmologist is to persuade practical astronomers that $D$ is not the only possible definition of the distance of $G$. If for instance, parallax measurements could be made on this galaxy, it is not difficult to prove that the distance so computed would not be identical with $D$ as soon as the red-shift attained 0.2 or 0.3. So deeply rooted are the ideas of Euclidean geometry and of Newtonian mechanics in the minds of most astronomers, that a statement such as I have just made is usually greeted with a polite but skeptical smile! The underlying theory of formula (1) reveals that $F(z)$ arises from the following two causes. Firstly because the photographic emulsion is not equally sensitive to light of all wave lengths so that shifting the entire spectrum towards the red distorts the impression made on the plate; secondly, the absolute magnitude of $G$ at the moment of emission of the light may be different from the absolute magnitude of a similar galaxy in the neighbourhood of $O$ with which $G$ is compared at the moment of arrival of its light at $O$. The first effect is known as the $K$-correction, the second as the Stebbins-Whitford correction. The use of formula (1) in the Neyman-Scott statistical theory requires care because integrations have to be performed over all positive values of $z$, whereas the form of $F(z)$ is known only for $z$ up to about 0.3. Admittedly, the contributions to the integrals when $z$ exceeds, say, one are likely to be small but for this, and many other reasons, an intensive study of the function $F$ is much to be desired.

It is sometimes argued that the red-shift versus apparent magnitude relation can be converted into a velocity-distance formula and that this gives the anwers to all distance problems in the universe. If this were so, it might not be necessary to use two kinds of distances in the Neyman-Scott theory, the space-interval in $f$ and luminosity-distance in $\theta$. But when we examine the velocity-distance formula, we find that velocity is defined by $cz$ and distance by $D$. Now it can be proved that, except for very small values of $z$, it is impossible to indicate in an unambiguous manner what distance it is whose rate of change is equal to $cz$. It is certainly neither luminosity-distance, nor distance measured by parallax, nor by apparent size, nor by any other method for which a clear operational procedure can be given. Velocity to my mind is a meaningless term unless the distance whose rate of change it is is definable, and therefore I regard the velocity-distance relation as a bastard formula in which velocity is defined through the classical Doppler theory in Euclidean space, and the use of absolute Newtonian time, and distance by a relativistic formula for luminosity-distance. Worse still, the velocity-distance relation is then used in relativistic cosmology where neither Euclidean space nor absolute Newtonian time has any meaning. To avoid these inconsistencies I would make a plea for a return to the observational data and to the empirical formula that expresses log $z$ as a polynomial in the apparent magnitude $m$, and in which neither distances nor velocities appear.

I have tried to give a general view, omitting all mathematical proofs, of the way in which the Neyman-Scott theory can be harmonised with relativistic cosmology. Much that I have said may have sounded like dogmatic and improbable assertion; if so, I would ask you to consult a paper [6] in which I have given the necessary proofs and, whilst reading it, I urge you once again to reflect that the scientific prejudices instilled in your minds by long acquaintance with classical mechanics are neither unalterable nor the expressions of self-evident truths.

## REFERENCES

[1] J. H. Bigay, "Photometrie photographique des nebuleuses extragalactiques," *Publ. Obs. Lyon*, Vol. 5, Ser. I, Fasc. 20 (c. 1952); *Annales d'Astrophysique*, Vol. 14 (1951), pp. 319–382.
[2] M. L. Humason, private communication.
[3] C. D. Shane and C. A. Wirtanen, "The distribution of extragalactic nebulae, *Astr. Jour.*, Vol. 59 (1954), pp. 285–304.
[4] J. Neyman and E. L. Scott, "A theory of the spatial distribution of galaxies," *Astrophysical Jour.*, Vol. 116 (1952), pp. 144–163.
[5] J. Neyman, E. L. Scott and C. D. Shane, "On the spatial distribution of galaxies: a specific model," *Astrophysical Jour.*, Vol. 117 (1953), pp. 92–133.
[6] G. C. McVittie, "Relativity and the statistical theory of the distribution of galaxies. *Astr. Jour.*, Vol. 60 (1955), pp. 105–115.