INDIAN STATISTICAL INSTITUTE

SIXTEENTH CONVOCATION ADDRESS

Statistical Techniques in Astronomy

by

J. V. Narlikar

Professor and Head, Theoretical Astrophysics Group
Tata Institute of Fundamental Research, Bombay

6th February 1982

203 BARRACKPORE TRUNK ROAD
CALCUTTA 700 035

Mr. President, Mr. Chairman, Professor Adhikari, Members of the Indain Statistical Institute and distinguished guests,

I consider it a great honour to be invited to address the Sixteenth Convocation of the Indian Statistical Institute. Indeed, may I confess to a feeling of awe which comes to me when I look at the list of distinguished persons who have preceded me ? Nevertheless I am grateful for this opportunity which enables me to highlight the impact your subject, statistics, has made on my subject, astronomy.

Statistics and astronomy

You may ask 'what can a hard headed statistician offer to a starry eyed astronomer ?' The answer is, 'plenty'. One normally associates statistics with large numbers, and astronomy is full of large numbers. The number of stars in our galaxy, the so called Milky Way System, is more than a hundred thousand million. The number of galaxies in the observable universe is upwards of a thousand million. Surely these large numbers justify a *prima-facie* case for the use of statistical techniques !

Indeed astronomers of various disciplines have compiled catalogues of various objects, stars, galaxies, radio sources, quasars, X-ray sources and so on. These catalogues have been subjected to routine statistical techniques. It is not however this routine aspect that I want to talk about today. My topic concerns the way statistical arguments are being employed in generating and resolving astronomical controversy. Since I am speaking before experts, let me at once confess to the charges of 'personal bias' and 'selection effects' in the instances to be cited now.

The Hoyle-Ryle controversy

In the January of 1961, when I had barely started my career as a graduate student at Cambridge, there suddenly erupted what has now become well known as the controversy between Fred Hoyle and Martin Ryle. As a student of Hoyle I found myself a somewhat bewildered participant in the hot arguments that ensued. Let me briefly outline the basic astronomical point at issue and then describe how statistical concepts became important.

The argument concerned the steady state theory of the universe, a theory which was proposed in 1948 by Hoyle and two other astronomers, Hermann

Bondi and Thomas Gold. As the adjective 'steady' implies, the physical properties of the universe do not change with epoch according to this model. Since by surveying the more remote parts of the universe we see it as it was in the past (—remember that our means of survey, the light wave, travels with a finite velocity—) the steady state hypothesis can in principle be tested by astronomical observations.

Martin Ryle proposed to do so by counting the radio sources which lie well outside our galaxy. If these sources are distributed uniformly, and if all of them are equally bright, then the following two conclusions can be drawn about their counts.

(I) The number N of sources out to a distance R should vary as the cube of R. (This is simply the Euclidean relation between volume and radius.)

(II) The faintest of the sources in the above set would have its flux density S (which is the radio astronomer's measure of how bright the source looks from here) varying as the inverse square of the distance R.

Combining (I) and (II) we arrive at the following conclusion :

(III) The number N varies with S as the inverse 1.5 power of S.

We therefore expect the number of sources brighter than a given flux limit to increase according to the above inverse 1.5-law as the flux level of the survey is progressively lowered. Ryle and his collaborators who conducted the Cambridge survey of radio sources did find N to rise as S was lowered. However, the observed rise of N was *steeper* than that predicted by the Euclidean 1.5-law just mentioned.

On this basis Ryle concluded that the fainter radio sources were more numerous and hence that there were more radio sources in the past than are seen now. Clearly, this finding contradicted the 'steady state' postulate. Further, the larger density of radio sources in the past implied, according to Ryle, a denser universe in the past and this idea was in support of the rival 'big bang' theory, according to which the whole universe is supposed to have originated in a gigantic explosion from a highly dense state.

Faced with this challenge, Hoyle set about to counter the above claim. I recall having to work very fast with him on a scenario which was not only consistent with the steady state postulate but which also reproduced Ryle's observations. Fast, because we had to produce our answer at the time that Ryle made his results public. As it happened, Hoyle had to be away on a prior engagement at the time that Ryle was to present his work at the Royal Astronomical Society. And so he delegated to me the task of presenting our counter-arguments.

Basically, our rejoinder rested on the premise that although the universe is steady and homogeneous on the large scale, it may have inhomogeneities on a small scale. Since the radio sources are not so numerous as galaxies, their counting is subject to statistical fluctuations. We showed that these fluctuations can cause significant deviations from the 1.5-law and that Ryle's steep curve could well arise from such an effect. The fact that such inhomogeneities like clustering and superclustering of galaxies exist in the universe was not so well known in the early-sixties as it is today.

In subsequent work I ran a number of hypothetical Monte-Carlo models on a computer to simulate clustering of various sizes. It became clear that hypothetical Ryle-type observers, looking from different locations would come up with the N-S relations both steeper and flatter than the Euclidean 1.5-law. A single survey from one vantage point (like the Cambridge survey) could no more be taken to disprove the underlying theory than a coin could be classified as biassed after a single throw.

The moral that emerged from this episode is that statistical fluctuations from the expected result can be significant if we are looking at a comparatively rare object.

The maximum likelihood method

As radio astronomy progressed from its early days in the mid-1950s, astronomers looked for better reduction of data and for more clearcut conclusions. It is here that statistics made important contributions. Let me give a couple of examples.

The first example again concerns the counting of radio sources down to varying levels of flux density. Suppose we have a plot of N against S and we expect a relation of the following kind to emerge

$$N \propto S^{-\alpha} \qquad \text{.. \qquad ..} \qquad (1)$$

where α is positive number. We have already seen that in a Euclidean universe with a uniform distribution of radio sources, all of equal intrinsic brightness, the above index α has the value 1.5.

Instead of having a previous notion about α, suppose the astronomer wants to know that value of α gives the best fit to the observed (N, S) values. In this connection a radio astronomer D. L. Jauncey pointed out that it was incorrect to take the data in the form of (N, S) values. Jauncey's argument is as follows: Suppose we have two sets of values, (N_1, S_1) and (N_2, S_2) with $S_1 > S_2$. Then by our definition, all sources which are included in N_1 are also part of the larger number N_2. In other words, N_1 and N_2 are not statistically independent. In going from S_1 to S_2 we should consider only the *additional* sources $N_2 - N_1$ as being independent of those previously counted at the level S_1.

Thus the index α is properly estimated by considering differential source counts, with the relation (1) changing to

$$\Delta N \propto S^{-(\alpha+1)} \Delta S. \qquad \ldots \qquad \ldots \qquad (2)$$

The estimate of α through differential source counts would therefore be more reliable than the estimate from the data on integral source counts.

Having pointed out this fact, Jauncey went on to describe a variant of the method known to the statisticians as the 'maximum likelihood method'. Using this method it is possible not only to estimate α but also to assign error bars to the estimate. Thus questions like whether the observed value of say, $\alpha = 1.8$ is really different from the predicted value of $\alpha = 1.5$ can be answered with appropriate levels of confidence.

Angular sizes of radio sources

My second example takes me to the important work on angular sizes of radio sources, which recently came out of the 550-metre radio telescope at Ooty. The Ooty telescope is admirably suited for measuring the angular size of a radio source as it is occulted by the moon. Before describing this work let me set the historical background to it.

In 1958 Fred Hoyle had pointed out a peculiar effect which arises from the non-Euclidean geometry of the expanding universe. The effect may be best seen in contrast to our day to day experience which is based on the geometry of Euclid.

A distant mountain looks smaller than a low hillock nearby. A real train viewed from a distance looks like a toy train. Our assessment of the physical size of an object depends on the angle it subtends at our eye. If we have a population of objects of the same physical size located at varying distances from us, the more distant ones will appear smaller than those closer to us. In fact as its distance from us tends to infinity, the apparent size of the object tends to zero as the reciprocal of the distance.

This observation is, however, based on Euclidean geometry ! Something, rather striking emerges in the corresponding case of the expanding universe. Imagine objects of the same physical size distributed all over the universe and viewed by us. What do we expect to find ? According to the calculations of Hoyle, the apparent size should start decreasing as we survey more and more remote objects. However, the rate of drop in size slows down and beyond a critical distance it eventually gets reversed, with the results that very far away objects would look bigger and bigger the farther are they from us ! Can this unexpected observation be checked in practice ?

Hoyle suggested that extragalactic radio sources be used for this test. Do radio sources at increasing distances begin to appear with increasing angular size as we see them beyond the critical distance ?

Two difficulties present themselves in the implementation of this programme. First, the astronomer does not yet know how to measure the distance of a radio source. He infers it from how faint it 'looks' : a remote source will be fainter than the nearby one according to the inverse square law of illumination as adapted to the non-Euclidean geometry of space and time. This method would work reliably if all radio sources were equally powerful. In practice the powers of radio sources range over at least five orders of magnitude, thus resulting in large uncertainties of distance. To put it bluntly : can we tell a strong but remote source from a weak but nearby source ? The honest answer to this question is "no". The astronomer therefore plots the angular size against the observed flux density rather than the distance of the source.

The second difficulty comes from the fact that radio sources are not spherical like stars. They are linear in structure, shaped more like dumb bells. Radio emission comes from the ends of the typical dumb bell shaped source. Now the source would look very small if it is aligned almost with the line of sight while it would look large if it is projected across the line of sight. Since the astronomer can measure only the projection of the source perpendicular to the line of sight he has to allow for an unknown projection angle to correctly estimate the 'true' length of the source.

Both these uncertainties necessarily introduce enormous scatter in the observed angular size flux density relation, apart from the recognized fact that, source for source, not all objects surveyed are of equal size. The intrinsic size variation is also over such a large factor as about 100 — 1000.

The median test

It was not surprising therefore that the radio survey from Ooty by Professor Govind Swarup and Dr. Vijay Kapahi produced a large scatter in the angular size-flux density plot of radio galaxies. To extract a signal from this noise they looked for the variation of the *median* angular size with flux density. Thus, they divided the flux density into a large number of bins and calculated the median of the angular size distribution in each bin.

This median size, however, did not show the rise at very low flux densities. That is, the phenomenon of increase of angular size with distance was *not* observed. To reconcile their finding with the expected result of Hoyle, Swarup and Kapahi had to assume that the more remote sources were systematically smaller in size compared to the nearby ones. Since the more remote sources are seen as they were at epochs much earlier than the present epoch, this conclusion implies a steady increase in the size of the radio sources with age—a conclusion in conflict with the steady state theory of the universe.

While this conclusion seemed reasonable for the median curve, it left the nagging feeling that perhaps too strong a deduction was being made from highly scattered data. As statisticians you are well aware of the fact that the trustworthiness of the average depends on how large is the spread around it. My colleague Dr. Chitre and I were therefore interested to find out an estimate of the reliability of the median curve. This is where statistics again came to our aid.

While there is considerable literature on significance tests on the mean, there is hardly any discussion of the median in any of the statistical texts commonly used by scientists. After some search, however, we did find a discussion of a median test in the book 'Theory of Rank Tests' by J. Hajek and Z. Sidak, which suited our purpose. Using this test we were able to demonstrate that the apparent difference between the observed median curve and that which could arise in a steady state universe was well within the 95% probability limits set by the test. Thus the scatter in the data is large enough to accommodate the steady state model of the universe.

Like the above example, there is another instance where astronomers are looking to guidance from statisticians for resolving an important question. Simply stated the question is as follows : "Are quasars local or distant ?"

Again, to set the background to this controversy, let me first say something about quasars. These remarkable objects were first identified by astronomers early in 1963 and by now more than 1500 quasars are known and catalogued. On the astronomer's photographic plate these objects appear as star-like images. Indeed the first two quasars to be discovered were mistaken for stars initially.

The most important difference between ordinary stars and quasars is the property of redshift. The spectrum of a typical quasar shows a systematic increase in the wavelengths of all familiar lines. The visible part of the spectrum is thus shifted towards the 'red' end. The fractional increase in the wavelength of a spectral line is called the redshift.

The phenomenon of redshift is not unknown for stellar spectra. In most cases the redshift in a star's spectrum arises from its motion away from us. The cause of the redshift is thus the well known *Doppler* effect and the magnitude of the redshift can be directly related to the velocity of the star. In a few stars the redshift can arise from another reason. In a compact star such as a white dwarf the strong gravitational field at the surface of the star causes light to be redshifted. This effect, known as the *gravitational redshift effect*, was first emphasized in Einstein's general theory of relativity.

In either case the amount of redshift observed in a star's spectrum does not exceed 0.001. The first two quasars to be discovered, had redshifts 0.158 and 0.367. These were considerably in excess of the redshifts found in the spectra of stars. The astronomers therefore felt inclined towards a third type of explanation for the quasar redshifts.

This explanation had however, been used earlier for galaxies, not stars. As early as in 1929, Edwin Hubble had reported redshift as a general feature of galaxy spectra. In fact Hubble was able to arrive at an empirical law according to which the redshift of a galaxy is proportional to its distance from us. At the time the quasars began to be discovered. Hubble's law for galaxies had become well established and the largest known redshift of a galaxy was 0.46.

It was natural for most astronomers therefore to adopt this interpretation for the redshift of a quasar.

This interpretation meant, however, that quasars are considerably farther away than stars in our galaxy. Whereas stellar distances range up to tens of thousands of light years, the quasars must be located at distances of hundreds of millions of light years. This also meant that although star-like in appearance, a quasar should be as powerful a source of energy as a galaxy (which contains upwards of a hundred thousand million stars).

A number of technical problems arise from the above conclusion, and astronomers have been attempting to come to grips with them. The main difficulty arises in thinking of a viable source of energy which is vast in output, yet compact in size. Some astronomers by mid-1960s thought these difficulties intractable and felt that the Hubble law may not, after all supply the correct explanation for the redshift of a quasar.

In modern cosmology, Hubble's law is deduced as a consequence of the expansion of the universe. For this reason the hypothesis that the redshift of a quasar obeys Hubble's law is known as the *cosmological hypothesis.*

Anomalous redshifts

To an outsider it would appear simple matter to decide whether quasars follow Hubble's law : measure the redshift and the distance of a quasar and see if the two satisfy the Hubble relation. The trouble is, we have no means of measuring the distance to a quasar and thus test the validity of the cosmological hypothesis. This is the crux of the controversy about quasars that I mentioned earlier.

Recently observations of H. C. Arp have added fuel to this controversy. Briefly, what Arp has done is to locate a number of high redshift quasars near galaxies of low redshifts. To understand what this type of observation implies let us take a hypothetical example.

Suppose we see a quasar Q of redshift 2 near a galaxy G of redshift 0.02. If Hubble's law applies to both the quasar and the galaxy we would conclude that Q is considerably (that is by a factor of 100 or so) farther away from us than G is. Under such a circumstance the observed nearness of Q and G on the sky is accidental ; their directions happen to be very nearly the same.

However, while galaxies are fairly numerous, quasars are not. Thus the distribution of quasars on the celestial sphere is like the distribution of a number of points with largish gaps in between. Under these circumstances, the chance of finding a quasar within a small angular distance from a given galaxy could be rather small. If the probability for such an event turns out to be extremely small, say, one part in a thousand, then by the conventional statistical standards, we would be entitled to doubt the validity of the whole chain of arguments based on the cosmological hypothesis.

This is what Arp does. He computes the probabilities for his observed quasar-galaxy associations and finds them to be very small. He therefore concludes that the quasars could not have been seen near the galaxies by a purely random process of juxtaposition of directions. From this conclusion he goes on to argue that the quasars and their respective galaxies are physically associated. That is, they are in the same part of space.

Such a conclusion vitiates the cosmological hypothesis. It implies that even with its large redshift the quasar could be a nearby object.

Naturally, the proponents of the cosmological hypothesis are up in arms. They criticize the work of Arp on several grounds, e.g., that he does not define his selection procedure, that he computes the probabilities wrongly and so on.

It is here that we may hope for some input from the professional statistician. Just how should one define one's sample in order to arrive at a reasonably accurate conclusion ? What statistical tests can meaningfully tell us that a distribution of quasars near a distribution of galaxies is non-random ? In what way can we estimate probabilities for the observed distributions to arise by pure chance ?

In this connection I may describe a few striking instances of the configurations observed. In quite a few cases Arp and his colleagues find a galaxy with redshift less than 0.1 surrounded by quasars of much higher redshifts. The quasar redshifts may range up to 3 ; but there is another striking feature. We find quasars of nearly equal redshifts on opposite sides of the galaxy ! Personally I find it very hard to imagine that such configurations arose out of random projections on the celestial sphere.

In 1980, Arp and Cyril Hazard noticed another peculiar configuration on a photographic plate. They found two triplets of quasars. The three quasars in each set are exactly aligned in a striaght-line, the observational accuracy of locating the quasar images being about 2 arc seconds. The redshifts in one triplet are 0.51, 2.15 and 1.72 while in the other they are 0.54, 2.12 and 1.61. To make the matters 'worse' two more such aligned triplets of quasars were located on the same plate by another astronomer, W.C. Saslaw. The probability of four such triplets arising purely by chance is estimated to be less than one part in 10,000.

This latter example has roused considerable interest. It can be argued, I think quite legitimately, that the *a-posteriori* calculation of probabilities can be misleading. We have to avoid Bertrand's paradox in sorting out what we mean by 'randomness'. Further, where probabilities are small, it may be dangerous to rely on a single plate for drawing conclusions.

The alignment problem has been studied recently by computer simulations of randomly distributed points on imaginary photographic plates. Here also there is a difference of opinion as to whether the type of configuration just described arises by chance with moderate probabilities. M. G. Edmunds and G. H. George think 'yes it can' while some recent work by K. Subramanian and myself gives a negative answer to the same question.

Is new physics needed?

The quasar problem is important because if Arp-type observations hold out the cosmological hypothesis is doomed. If quasars are local then what can their high redshift be due to ? Should we fall back on the other two explanations, viz., the Doppler redshift and the gravitational redshift ? My own recent investigations suggest that there may be problems with either of these alternatives, although the problems may not be insurmountable.

If these alternatives don't work we may have to look for a fourth interpretation for the quasar redshift. In 1980 P. K. Das and I suggested such an interpretation based on the gravitation theory of Fred Hoyle and myself. However, such an interpretation would be resorted to only if the other, more conventional alternatives fail.

This brings me to my concluding remarks. Astronomy, by its very nature forces the physicist to think about conditions far more extreme than seen and tested in the terrestrial laboratory. There are astrophysical phenomena which propel us to the frontiers of existing knowledge, to a domain where facts and speculations merge. The tendency of the conservative physicist is to hold on to the well established laws and view any non-conforming data with suspicion. The radical physicist on the other hand may like to think the data to be important enough to warrant a new law of physics.

While a number of personal factors enter the arguments between the two sides, I feel that statistical techniques can be fruitfully used to introduce objectivity. The examples I have cited during the course of this address show how useful statistics can be to astronomy. Conversely, by posing new problems astronomy may inspire new techniques in statistics. With valuable data to come in future from such sophisticated instruments as the Space Telescope, I have every reason to believe that increased interaction between statistics and astronomy will be to the benefit of both the subjects.

Let me therefore thank you once again for giving me this opportunity to air my views before you. I also take the opportunity to congratulate and extend my good wishes to those who received their degrees and awards today. I am confident that your distinguished institution will continue to provide experts capable of making valuable contributions in the pure as well as applied sciences.