

UC Irvine

UC Irvine Previously Published Works

Title

Statistical Challenge in Medieval (and Later) Astronomy

Permalink

<https://escholarship.org/uc/item/6165d8xh>

Author

Trimble, Virginia

Publication Date

2003

DOI

10.1007/0-387-21529-8_1

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

1

Statistical Challenge in Medieval (and Later) Astronomy

Virginia Trimble¹

ABSTRACT Portions of the history of the interaction between astronomy and statistics are told in the form of short case studies of a number of people who appear (or should appear) in books about both. These should be regarded as notes for a serious discussion of the subject, not the discussion itself.

In memory of Peter August Georg Scheuer from whom I (and many others) first heard that $N^{\frac{1}{2}}$ is sometimes signal rather than noise.

1.1 A demographic introduction

If one is going to explore the contributions of astronomers to statistics and of statisticians to astronomy, one ought perhaps to begin by deciding what is meant by an astronomer, a statistician, and statistics. I will not do so, and merely call attention to the cases, first, of Roger Boscovich of Dubrovnik, who rates a whole section in Hald (1998) for extending the method of least absolute deviations beyond where it had been left by Galileo for application to astronomical observations of latitude but is known only to the subsets of astronomers who collect foreign paper money or speak Serbo-Croatian (in which his name is spelled - and pronounced - quite differently) and, second, of John Michell, who appears in lots of astronomy treatises for inventing the concept of black holes, and, occasionally, for the discovery of binary stars, but does not make it into the statistics histories of Hald (1998), Stigler (1986), or Pearson (1976), despite his binary task having been accomplished by a method that most of us would call both statistical (certainly probabilistic) and innovative.

How large is the overlap between the two communities? Of the 76 astronomers indexed in Abell (1982) who flourished from ancient times up to

¹Department of Physics and Astronomy, University of California, Irvine, CA 92697, and Astronomy Department, University of Maryland, College Park, MD 20742

about 1850, 49 (from Airy to Zach) appear in one or more of the statistical histories by Stigler (1986, 1999), Hald (1990, 1998), Pearson (1976) and Franklin (2001). They, in turn, mention another 27 astronomers (Arago to Thomas Young) who did not make the Abell cut but who are mentioned in Russell, Dugan, and Stewart (1926), in Hoskin (1999) or some other reasonable place. RDS was the primary introduction astronomy text in English for about 20 years. George Abell wrote the first of the now-ubiquitous books for non-science major courses, with the 4th, 1982, edition the last over which he had control. And Hoskin's volume is the most recent attempt to put the entire history of astronomy between two covers.

Closer to the present, scientists become more and more specialized but in the period from 1850 to 1950 at least the following can reasonably be described as having contributed to the astronomy/statistics interface: Simon Newcomb (1886), Arthur Eddington (1914), and Harold Jeffreys (1939), noted by Hampel (1998), who also regards the work of Cannon, Fleming, and Leavitt as statistical in nature, Jacobus Kapteyn (1922, ending his 40+ years of work on the topic), Jerzy Neyman and Elizabeth Scott (1956), W.M. Smart (...), S. Chandrasekhar (1939), Gunnar Malmquist (1920, 1924), Col. Frank J.M. Stratton, and Robert Trumpler and Harold Weaver (1953).

The sign of the contribution is not always clear. Consider the case of Stratton, who was the last person to have participated officially in every general assembly of the International Astronomical Union and who was one of the officers who held the Union together during the very difficult 1939-1945 period, but whose astronomical work most of us would be hard-pressed to recall. He was also the Cambridge tutor of Ronald Fisher (of the F-distribution and much else), and I cannot resist quoting the following from Hald (1998):

The astronomer F.J.M. Stratton (1881-1960), who was Fisher's tutor, lectured occasionally on the theory of errors. We do not know precisely the contents of his course, but in the preface to a book by D. Brunt (1917), the author thanks Stratton, "to whose University lectures I owe most of my knowledge of the subjects discussed in this book, and upon whose notes I have drawn freely." There is nothing original in this book.

Not knowing Hald, I cannot be sure whether he means this to be as mirth-provoking as it is. Stigler (1999), on the other hand, clearly means to amuse as well as to enlighten when he includes in a section called "Questions to Discovery" a chapter entitled "Who discovered Bayes' Theorem?", one called "Daniel Bernoulli, Leonhard Euler, and Maximum Likelihood" (to which a local wit responded, "Oh, yeah. Old Max. He used to drink a lot."), and one called "Gauss and the invention of least squares." The issue of which items in astro-statistics and statistico-astronomy should be called

discoveries and which inventions is another issue that I will not resolve here. Indeed I will say nothing about Gauss and least squares, since his contributions, the antecedents, and descendants were so well explored by Rao (1997) in SCMA II.

What will appear in the rest of this paper is a series of case studies, of what strike me as fruitful interactions between the fields. None is precisely medieval (how sure are we that the number of cardinal sins falls between 4.65 and 9.35?), though some archeoastronomy items appear at the end. Just how many of the tales get told will depend on the editor, who will remove as many as necessary to get below the assigned page limit.

1.2 Giants in the Land

These stories concern scientists of enormous reputation over a range of disciplines, and I have not consulted the original literature, but retell from Franklin (2001), Hald (1990, 1998), Hoskin (1999), Stigler (1986), Pearson (1976), and other sources read too long ago to be honorably recalled.

1.2.1 *Galileo and Least Absolute Deviations*

In the simplified version of history we hand our students while they are getting settled into their seats at the beginning of a lecture, the Aristotelian-Aquinian principle of “the immutability of the heavens” was overthrown by Tycho Brahe (1546-1601), who set an upper limit to the geocentric parallax of his nova stellar of 1572 (and also the comet of 1577) placing them beyond the sphere of the moon. But, not surprisingly, he was not the only astronomer of the time to look for this parallax. Incidentally, seeing the geometry of it is rather tricky for modern eyes, but it is a true statement that the new star, if it is close to the earth and turning in the diurnal motion about the pole, will show itself more distant from the pole when it is below the pole on the meridian than when above it (roughly Galileo’s words). A certain Scipione Chiaramonti (1565-1652) combined some of Tycho’s observations with those of 11 other astronomers to conclude that what we now call NS 1572 was at most 32 earth radii away, with similar conclusions for SB 1604 and the nova stella of 1600 (actually Mira).

This provoked Galileo (1564-1642) in his 1632 **Dialogo** to look again at all the reported measurements of upper and lower culmination altitudes of the 1572 star made by astronomers at latitudes from 38.5 to 56° north. That is, he is looking for geocentric parallax over a fairly small baseline rather than for earth rotation parallax which can be measured by a single observer and, for circumpolar stars (as SN 1572 was for Tycho) has a baseline of $2 R_e \cos(\text{latitude})$.

Galileo then compared the sums of the absolute values of the errors of the

observations implied if the distance was $32 R_e$ vs. sufficiently large to yield no parallax. Of the more than 100 pairings of the data points available, Galileo picked 10 most favorable to Chiaramonti's hypothesis and 10 most favorable to his (with no overlap). The sums of the absolute errors in the two cases were 756.9 arc minutes and 83.7 arc minutes respectively. Small being good in this context, he regarded the result as being strong evidence for a translunar location for the event. And so do we.

The method then languished until 1755, when Boscovich applied it to the determination of the lengths of arcs of latitude at various locations (in connection with the problem of determining the shape of the earth - prolate had been claimed). Galileo was also the first to figure out the odds of getting various outcomes from the case of three dice. I checked his numbers by writing down all the combinations, which is presumably how he did it. He got it right, and it is left as an exercise for the Gosset to figure out how the results would change in the case of fermionic or bosonic, rather than distinguishable, dice.

1.2.2 *Edmond Halley and survival rates*

Halley (1656-1742) is known to astronomers best for his prediction of the return of the period comet now bearing his name. On the astronomical side, he also discovered proper motions of the stars and secular acceleration of the moon, accurately predicted the path of the eclipse of 1715 over England, and served as Astronomer Royal from 1720 until his death (succeeding Flamsteed, who was first).

But he also wrote, in 1694, "... on... degrees of mortality... and prices of annuities." The end of the title makes clear why men of practical bent were concerned with human survival and death rates as a function of age. His work in this area is an interesting illustration of what our grandmothers called "making do with what you have." Since it was English annuities for which he was trying to set a fair price (or anyhow one that people would pay and that would not bankrupt the issuers), he would obviously have liked to have rates of the deaths of English persons (not just men, since annuities were often purchases for widows) as a function of age. But the methods of recording births and deaths in England, mostly in parish registers, did not provide the numbers needed, so he used tables of numbers of births and deaths and total population for Breslau.

According to Pearson (1976), Halley was probably also the author of a 1699 piece in **Philosophical Transactions of the Royal Society** (the first scientific periodical in any language, in case you wondered) called "a calculation of the credibility of human testimony." This is also phrased in the language of how much you should be willing to pay for things. For instance, of someone who is 90% reliable tells you that he has seen your cargo ship safely into the harbor and unloaded without damage, then you should be prepared to pay (only) 10% of the value of the cargo to insure against

the loss of the whole. The paper does not address how you determine the reliability quotient of your informant, which is the aspect of the problem on which we most often stumble even today, whether the issue is astronomical or financial.

1.2.3 *Tobias Mayer and the libration of the Moon*

Mayer (1723-1762) tackled a problem whose geometry is even more difficult to see than that of geocentric parallax and solved it, using a method (called Mayer's or, more often, the method of averages) that would elude Euler working on the mathematically rather similar problem of mutual perturbations in the orbital motion of Jupiter and Saturn. Mayer's goal (connected with the use of lunar motion for longitude determination) was to find three angles: the one between the true rotation axis of the Moon and the poles of the circumference parallel to the ecliptic, the ecliptic longitude of the node at which the plane of the lunar rotation equator crosses the ecliptic, and the true latitude of the crater (Manilius, a suitable choice in several ways) he had observed. The observations were 27 pairs of angular positions of the crater parallel and perpendicular to (changing) apparent equator of the moon (the circumference parallel to the ecliptic), gathered by him over a couple of years.

Thus he had 27 equations in three unknown. His solution was to group these in three sets, with large, medium, and small (negative) coefficients of the first angle mentioned above, which he regarded as the most important. He then added up the groups (he could alternatively have averaged them) and solved the resulting triple, concluding that the result would be more accurate than that from any three data pairs alone (true) by a factor nine (false; it is at best three if only random errors in the observations are important). He apparently invented \pm as well.

Euler, writing in 1749 (the year before Mayer) was faced with 75 sets of observations of Saturn and Jupiter, gathered over 163 years, from which to extract eight unknown describing the orbits and their interactions. He pulled out the two that were not periodic in the 59-year synodic period of the two planets, and then ground to a halt, when various combinations of the equations led to wildly inconsistent results, saying that the errors had multiplied themselves through combining of observations. Nevertheless, most of us have heard of Euler, and few of Mayer. Indeed, Stigler (1986) notes that the method of averaging (or summing) equations discovered by Mayer is often attributed to Euler. His section heading is, of course, **Saturn, Jupiter, and Euler**.

1.3 Three careful clerics

James Bradley (1693-1762, third Astronomer Royal) and John Michell (c. 1724-1793) turn up in historical astronomy discussions with the words statistics or statistical attached to their persons. Bradley is known to Stigler, but not to Hald, and Michell to neither. These are the two stories for which I returned to the original literature and both remain high on my list of favorites, even after reading many pages in which f is pronounced s .

1.3.1 *Bradley and the aberration of starlight*

Bradley set out to find (as many others before him, and after, did) heliocentric parallax as the definitive demonstration of Copernican cosmography. He focused initially on Gamma Draconis, chosen by Robert Hooke for the same purpose, because it comes very close to the London zenith, thus minimizing both atmospheric refraction and flexure of the observing apparatus. By great good luck, the star is also very close to the ecliptic pole. The critical papers are Bradley (1728 on nutation). Hirshfeld recounts many more details than there is space for here.

Aberration is the apparent shift in positions of all stars (independent of distance) caused by earth's orbital motion. The maximum displacement is the ratio of orbit speed of light (10^{-4} or a half-angle close to 20 arcseconds), and the standard analogy is walking forward into falling rain and needing to tip your umbrella to keep the drops from hitting you. Bradley seems to have found geometry easy and does not sketch the situation. Incidentally, he is able to report observations taken right through the year. You cannot see stars by daylight from the bottom of a well, but you can with a suitable (preferably long local-length) telescope.

Aberration shows in a year (or less) of data as our direction of motion through space changes and a star near the ecliptic pole seems to move north and south in declination at transit. Bradley continued to follow Gamma Dra over the years at the same time as he moved on to other stars, seeking to confirm the effect. After 20 years, it became clear that there was a systematic residual, with period about 19 years, which we now call nutation and attribute to lunar tides. His second paper makes use of (at least) the following ideas that are statistical in nature:

- (a) mean values for the rate of precession of the equinoxes and obliquity of the ecliptic (rather than a favorite, or the most recent, or the oldest);
- (b) a weighted mean for the maximum value of aberration for a star exactly at the ecliptic pole, which takes into account data on about 10 stars, giving largest weight to Gamma Dra, which has the longest data string, the smallest polar and ecliptic polar distances, and the brightest apparent magnitude; and

(c) an examination of the distribution of residuals.

He says that, in the comparison between observed declinations (or altitudes at upper culmination) and ones calculated from his final model, 11 of 300 values differ by 2-3" and none by more than 3".

Bradley ends by noting that he suspects that some physically meaningful effect remains to be found (e.g. a secular decrease in the obliquity). In modern terms, the fact that the distribution of errors is flatter than a Gaussian with a standard deviation of 1 arc second is confirmation of his suspicion. He displays a number of tables of observed and model declinations, one of whose implications is that, in 1748 in England, the autumnal equinox came about September 9th.

1.3.2 *John Michell and binary stars*

Michell is also part of the quest for parallax, because his demonstration that pairs of stars close on the sky are generally bound systems rather than chance superpositions undid the hopes of William Herschel and others to use such pairs for parallax measurement, on the assumption that the fainter star would always be more distant. He also, of course, thereby demonstrated that not all stars have the same absolute brightness, enormously complication "star gauging" or "the" problem of statistical astronomy (next section).

Michell appears in various contexts as:

- (a) the inventor of black holes ("all light emitted from such a body would be made to return to it, by its own proper gravity." Michell, 1783);
- (b) designer of the Cavendish balance (Cavendish was his executor),
- (c) propounder of the idea that earthquake energy travels in waves (based on times at which Lisbon 1755 shook up other European cities); and
- (d) the discoverer of binary stars (though it took Herschel's measurement of the first bit of an orbit before all were persuaded).

Michell (1767) began by asking for the probability that any one particular star should happen to be within a certain distance (as for example one degree) of any other given star and finding that it is $(60)^2/(6875.5)^2$ or $1/13131$. And the probability that it is not is $13130/13131$. He then extends to the probability that no one of whole number of stars n would be within one degree from the proposed star, and its complement, $1 - (13130/13131)^n$ that there is one, and so onward to the probability that no one star should be within a distance r of any other star, with n to choose from,

$$P(\text{not}) = \left(1 - \frac{r^2}{(6875.5)^2}\right)^{n \times n}$$

and its complement, the probability that one is.

He makes fairly heavy going of the arithmetic, ending up with a style that resembles that of a modern student whose calculator doesn't have quite enough significant figures in its chips. Apparently $(1+n)^x = 1 + nx + \dots$ was not part of the standard tool kit, but he gets the right answer, finding, for instance, that for Beta Cap ($n = 230$, $r = 3.3 \text{ arc min}$) the chances are 80:1 against its being a chance alignment. For the six brightest Pleiads, the odds are 496,000 : 1 against a chance grouping.

If this sort of arithmetic rings a bell, it is probably because you have met it before as the question of how many non-twins must you have in the room before it becomes more likely than not that two of them have the same birthday. The number (about 22) is smaller for Moslems because their year is shorter. I have no idea whether Michell or his predecessors knew about the birthday problem or other events described by the same calculation, but he does seem to have been first to apply it in astronomy.

1.3.3 *Nevil Maskelyne and the personal equation*

Maskelyne (1732-1811), the fifth Astronomer Royal, like Bradley and Michell, held orders in the Anglican church and is the member of the trio one finds it hardest to associate with the concept of charity, perhaps because he figures as something of a villain in the story of the quest to determine longitude at sea. He was indeed a supporter of the method using the motion of the Moon (Maskelyne, 1762), mentioned in connection with Mayer's work. He was also in some sense the discoverer of the first recognized systematic error in astronomy, generally known as the personal equation.

Back in 1796, when the right ascensions of stars were determined from their times of meridian transit, Maskelyne noticed that his assistant, David Kinnebrook, whose work had formerly been consistent with his own, was now recording transit times that were systematically 0.8 sec later than his own. This corresponds to 12 arc seconds or as much as 0.2 miles at sea, and this 68 years after Bradley had measured the polar distances of stars to 1 arc second or better. Rather than rejoicing in the discovery that systematic errors could be much larger than random ones (and that Bradley had been wise to measure altitudes rather than hour angles), Maskelyne waxed wroth and fired Kinnebrook. Twenty-some years later, Bessel (who eventually found the long-sought parallax) looked again at hour angles measured not only by Maskelyne and Kinnebrook but also ones of his own and from Struve (another parallax discoverer), Argelander, Walbeck, and Knorre and found systematic differences up to a second (of time) and more which could vary from year to year.

His way of writing these, as, for instance, $B - S = -0,799 \text{ sec}$. appears to have given rise to the name "personal equation" (Stigler, 1986). The magnitudes and variations were the sort normally associated with human reaction times, as per the story of Galileo's attempt to measure the speed of light

with dark lanterns on the seven hills of Rome. The name personal equation became customary and the numerical values dropped only with the adoption of automatic and electrified chronographs. The very large difference in **systematic** accuracies of right ascension (with personal equation) and declination data (without it) propagated through astronomy in the form of separate analyzes of the two components for many purposes, statistical and others.

The term was sometimes used for systematic errors of other sorts, for instance by Russell et al. (1926) to describe the tendency of some Mars observers to draw thin, straight lines between the dots and others to avoid this at all costs. Sherlock Holmes uses the phrase to mean something like general intelligence, remarking at one point that he “need not allow for what astronomers call the personal equation” since a particular foe is of first-rate intelligence (like himself, of course).

Any astronomer will be able to come up with other examples of unrecognized systematic errors utterly swamping the recognized random ones. Stigler’s (1999) Table 20.1 shows 15 successive published values for the length of the astronomical unit in miles. Only two fall within the error bars of the previous value, and only two have error bars that take in the present official number. This is known to 9 significant figures in metric units (from radar travel times), but only about 6 in miles (owing to disagreement about the conversion factor). My own favorite is the Hubble constant, which has declined from 536 km/sec/Mpc (Mpc = megaparsecs) according to Hubble’s initial, 1929, calibration, down to about 65, with 10% error bars at every stage (Trimble, 1996).

Maskelyne also makes a cameo appearance at the beginning of our next story, because he provided some of the key proper motion measurements from which Herschel first charted the motion of the sun relative to its neighbors. Other numbers came from Tobias Mayer, whom you have now also met.

1.4 “THE” problem of stellar statistics

Newton thought of, Michell (1767) and undoubtedly many others developed, and William Herschel is generally given credit for applying the method of determination of the distances and distributions of the stars in space based on the assumption that they are as bright as the sun (see Hoskin, 1963, for details of this story). Herschel called the method star gauging (gaging in his spelling) and by 1785 had put the sun near the center of a flattened system having sharp edges, a uniform density of stars, and an extent of a couple of kiloparsecs, stretching furthest in the directions where we see the most and faintest stars (“the Milky Way”). Even the Cygnus rift is there.

From Herschel's time down to the present, the key problem marching under the banner statistical astronomy has been to turn counts of numbers of stars as a function of apparent brightness into, in historical order:

- (a) the size and shape of the system;
- (b) the real distribution of stellar brightnesses (after Michell et al. showed that they were not all the same); and
- (c) the distribution of the velocities of the stars (as a function of location, brightness, and so forth) after proper motion data and, later, radial velocity measurements showed that the system is not a static one.

Trumpler and Weaver (1953) mark the high-point of this endeavor as a core subject in astronomy.

Why is it a statistical problem? The number of stars you count as a function of apparent magnitude, $A(m)$, is given by

$$A(m) = \omega \int_0^\infty \varphi(M)D(r)r^2 dr$$

where ω is the solid angle you are examining, $\varphi(M)$ is the luminosity distribution, and $D(r)$ is the density of stars as a function of distance in that direction. The implied assumption that $\varphi(M)$ and $D(r)$ are separable functions is already a fatal error if you propose to look more than about one kpc in the galactic plane or 100 pc perpendicular to it. Built in is the relation between apparent and absolute magnitude, $M = m + 5 - 5 \log r - a(r)$, where $a(r)$ is the absorption in magnitudes and constitutes another unknown function. Kapteyn (1922) was the last to do this for $a(r) = 0$ everywhere (though he had earlier suggested values of 0.3 and 2.0 mag/kpc in the galactic plane), and even in this case, one clearly has to go over to sums rather than integrals, leading to a Mayer- or Euler-like problem of many equations in many (but fewer) unknowns and the potential for ending up with nonsense through what Euler called the multiplying of errors (both Gaussian and Poissonian in this case).

McCuskey (1965) and van Rhijn (1965, Kapteyn's colleague and successor) summarize the additional computational difficulties introduced when $a(r) \neq 0$ and make it clear when the confirmation of spiral arms in the Milky Way was left for the radio astronomers (for whom $a(r)$ really is 0 most of the time).

Now try to do the dynamical (stellar population) problem, where the goal is to extract, for instance, $N(M, V)$ from observations of $A_1(m, \mu)$ and $A_2(m, V_r)$ in various directions in the sky, subject to the same unknown $a(r, \theta, \emptyset)$ and the non-separability of the luminosity function, the density distribution, and the kinematic properties. Apart from everything else, one simply must have the counts, apparent magnitudes, proper motions, and radial velocities for the same stars in the same directions in the

sky. Kapteyn's (1906) **Plan of Selected Areas** sought to address this problem. The IAU Commission (32) on Selected Areas eventually voted itself out of existence, but this is the one context in which Kapteyn's name is remembered today in a positive tone of voice. Binney and Tremaine (1987), the relatively modern authority, mention neither Kapteyn nor his star streams, but do make contact with his period via the velocity ellipsoid of Karl Schwarzschild (which has, among other things, the shape of a Gaussian normal in two or three dimensions).

"Data products" from the traditional endeavor called statistical astronomy include:

- (a) the luminosity distribution(s) of stars (which we now immediately try to turn into the mass distribution;
- (b) the solar motion (first found by Herschel, using proper motions from Mayer and Maskelyne); and
- (c) galactic dynamics.

The local distribution of stellar motions was described by Kapteyn as two star streams and by Schwarzschild as an ellipsoid. Neither means quite what you might guess, and I recommend Russell et al. (1926) or their references, Campbell (1913) and Eddington (1914) for clearer expositions than found in more modern references. All wrote before Trumpler (1929) forced galactic absorption upon us. Even so, the problem, in the words of RDS,

The problem of stellar statistics is to deduce from the apparent distribution of the stars in the heavens with respect to magnitude, proper motion, radial velocity, parallax, galactic concentration, etc. ... what is the true distribution of the stars in space ... in terms of three statistical functions: the density function, which gives the total number of stars per unit volume. ... the luminosity function, which shows what proportions of these stars have absolute magnitudes lying in successive equal intervals; and the velocity function, which defines the similar distribution of their velocities in space.

must be sung as "to invert the impossible matrix".

Against this background, the discovery of galactic rotation by Bertil Lindblad and Jan Oort might seem nothing less than miraculous. They did however, have the rotation of M81 (Max Wolf) and M31 (Vesto Melvin Slipher) to guide them.

1.5 A smattering of archeoastronomy

Archeoastronomy includes (at least) two territories - the use of ancient observations to shed light on current questions (Chinese and other sightings of comets and supernovae are the classic examples) and the use of modern astronomy to shed light on ancient cultures (the classic example is the Star of Bethlehem, which I shall not mention at all, statistical considerations not often being important for single events, whether or not miraculous, but this is perhaps as good a place as any to record my prejudice that Bayesian methods, while excellent for changing your mind by a small amount, are much less useful on the road to Damascus).

Was Ptolemy to be trusted? Two aspects of this question have a “goodness-of-fit” answer. First, it seems that some of his observations are “predicted” so well by his model that they must have been back-calculated. This “excessive goodness-of-fit” result is an old one (Newton, 1977). Second, very recently, Schaeffer (2001) has asked whether Ptolemy borrowed his catalog from Hipparchus, and, if so, did it leave a statistical trail. Because the two lived at different latitudes and in different centuries (with precession of the equinoxes), different stars skimmed their horizons with differing degrees of visibility (hence opportunities for accurate observations of position and apparent brightness). The conclusion is that his fourth-quadrant stars are borrowed, the first three new observations.

Alignments of pre-literate and peri-literate monuments have been scrutinized for astronomical significance from the time of Locker to the present (Krupp, 1988, is a good source.). Conclusion range from, “you can see the whole of positional astronomy, including precession and changes in the obliquity at Stonehenge” to “yeah, the door is on the north side.” I have dabbled in the now very densely populated part of this territory occupied by the pyramids of Giza (Trimble, 1964). Objectively, one can say things like

- (a) the inclination of the shafts from the King’s chamber of Cheops’ pyramid point (to the accuracy within which they can be determined) to the north celestial pole (where there was a star when the pyramids were built) and to the upper culmination of the middle star in Orion’s belt;
- (b) the main exit of the Great Temple of Amon-Ra at Karnak points northwest, but misses the direction of sunset at summer solstice by more than the accuracy of the measurement (1.0° at the time temple was built, Krupp, 1988); and
- (c) main axes of 38 other temples built during the Empire period point in 38 other directions, 7 close to the cardinal directions and 6 (NW), 7 (NE), 13 (SE), and 5 (SW) in each of the quadrants (Badawy, 1968, p.184).

You could ask a statistical question about how likely this is to be a chance distribution (and answer it by frequentist or Bayesian methods). But if the answer is to be a contribution to Archeoastronomy, then you must decide what hypothesis you are testing. The choices include perpendicular to the nearby riverbank or to the cliffs behind as well as astronomical orientations.

The next step is supposed to be to test the hypothesis against a new, independent data set, or, failing that, to attempt to multiply the chance probability you find (which is always very small or you wouldn't be doing this sort of thing) by the number of other hypotheses that would be equally interesting. In the Empire Temple case, there is no comparable sample, but lots of hypotheses, and you are left with the usual result, "well, maybe there is something there."

Section 3.3 carried the moral that systematic errors are nearly always larger than random ones. The lesson here is that you must choose a testable hypothesis and stick with it. "Part of Ptolemy's catalogue is more consistent with observations made from Hipparchus' 4-dimensional location than with observations from Ptolemy's own 4-location" is such a hypothesis. "The Egyptians deliberately lined up their temples and pyramids to incorporate astronomical information" is not. Investigations of non-cosmological redshifts (which are now more than 35 years old) seem to me also to suffer from a surfeit of shifting and untestable hypotheses.

1.6 Ancient statistics in modern astronomy

Recent forays of astronomers into statistical territory come sometimes perilously close to reinventing the wheel and making it square. Nevertheless, I think each of the following issues is still a live one and still on the interface.

Density of matter (including dark matter) in the galactic plane
This belongs to the tradition territory because the key equation is

$$\left[\frac{d}{dz} \ln \frac{N(z)}{N(z_0)} \right] < V_z^2 > = -4\pi G\sigma_0$$

where $< V_z^2 >$ is the component of the velocity ellipsoid perpendicular to the galactic plane and the logarithmic gradient is that of the density of stars perpendicular to the plane. The desired density is σ_0 , and the error made if you choose to take $\pi = 3$ will be smaller than other that are unavoidable. The equation and its application go back to Kapteyn and Jeans, though Oort often gets credit, and forward into modern models of the galaxy from Bahcall and Soneira, Kuijken and Gilmore, and others. The main errors are now recognized as systematic rather than random (though the latter are not small), because star populations change systematically away from the galactic plane, rendering color-based parallaxes too large (distance too small) because the more distant stars will be of lower metallicity, loser

mass, and more advanced evolutionary stage. Kapteyn and Jeans actually bracketed modern results, with $\sigma_0 = 0.099$ to $0.143 M_0/pc^3$, and we remain uncertain about whether there is a separate disk dark matter component.

Closely related is the attempt to estimate the contribution of very faint, low mass stars to the total density. Small scale surveys (like those from the Hubble Space Telescope) yield a handful of brown and old white dwarfs (random errors win), and large scale ones suffer calibration errors (one of which the late Willem Luyten ungenerously dubbed the Weistrop Watergate).

Malmquist bias and the Scott effect Wherever two or three cosmologists are gathered together, one will say that the others do not understand these: their essence, the difference between them, or how to correct for them. Adriaan Blaauw even objects to the term Malmquist bias, on the ground that the concepts are all to be found in earlier papers by Kapteyn.

$\log N - \log S$ This is a cumulative distribution of source numbers vs. apparent flux. Errors due to binning are thereby removed, but others introduced. Early applications in radio astronomy suffered from confusion (meaning two or more faint sources getting counted as a single brighter one), though the conclusion that there are more distance radio sources than nearby ones stands. Giacconi (1972) used it at a time when very few X-ray sources were yet known or identified to predict that the X-ray background would eventually be resolvable into many distant sources. He too was right.

$P(D)$ and $N^{1/2}/N$ The concept that Poissonian fluctuations in numbers of sources within your beam will translate into apparent fluctuations in background surface brightness has been rediscovered at every wavelength. Scheuer (1957) used it to add a few points to the $\log N - \log S$ curve from the Third Cambridge Catalogue (rousing the wrath of the then-powerful steady-state community). Applied to optical observations of elliptical galaxies, it is one of the newer subrungs on the distance ladder (because you can pull out the brightness of the individual brightest stars contribution, declare then to be on the red giant tip, and get a spectroscopic parallax). Applied to the X-ray background, the calculation shows that the number of sources needed is just about what you would get from a $\log N - \log S$ extrapolation, if the background is to be neither more ragged nor smoother than what we see (these sources have now been resolved by Chandra and other missions).

V/V_m was Maarten Schmidt's way of taking into account that he had a flux limited sample with both radio and optical flux limits so that he could use measured redshifts of a very small number to conclude that quasars were commoner in the past. He has said that the basic ideas can again be found in Kapteyn's work (Schmidt, 2000). Recently he has suggested (Schmidt, 2001, personal communication) that the same methodology applied to gamma ray bursters implies that those of short duration are closer (and less beamed) than those of long duration (which optical redshifts now exist).

The **Lutz-Kelker correction** is needed when you look at groups of measured parallaxes encumbered with measurement errors, which are intrinsically asymmetric (since no real parallax can be negative; Chiaramonti had trouble with this!).

Kaplan-Meier survival curves This is my own particular square wheel, honed when I was trying to figure out how to show (or anyhow display) data concerning the long-term publication records of astronomers starting out with Ph.D's from high and low prestige graduate schools. The principle end point was, therefore, ceasing to publish. But it seemed to me (Trimble, 1991) that posthumous publication was an unreasonable expectation (not true - Lundmark was co-author of a 1999 paper), and it removed the deceased from the set of those at risk, so that the curves could turn back up if more people in a cohort died than stopped publishing for other reasons.

Properties of binary star populations There are at least two issues. First, how do you allow for unresolved binaries when counting stars as a function of apparent brightness (part of "the" problem of stellar statistics). This cannot be dealt with until you know the answer to the second issue, what are the real distributions of binary periods, separations, mass ratios, eccentricities and all as a function of age, chemical composition, and whatever else matters. These all fold into various attempts to understand chemical, luminosity, and other evolution tracks for galaxies or their separate stellar populations. Much ink has been expended since Kuiper (1935) interpolated (correctly) and extrapolated (I thing incorrectly) from the handful each of visual and spectroscopic binary orbits available to him. I abandoned the fray in 1990, with the parting shot that the answer you get will depend on the sample you choose to look at. This remains true. Complete information could be obtained only by working to sharp limits in apparent magnitude, magnitude difference and separation (for visual binaries), velocity amplitude and period (for spectroscopic binaries), and light amplitude and period (for eclipsing binaries) and then carrying out the equivalent of V/V_m in about six-dimensional space to get a volume limited sample. This is (marginally) possible for nearby clusters. "all the F V stars in the Yale Bright Star Catalog" or a few other narrowly circumscribed classes, but otherwise impossible.

Can we derive any particular lesson from these more complicated cases? I think so (and it is one that spectroscopists working on stellar structure and evolution were forced, kicking and screaming, to accept a couple of decades ago). It is that, when comparing hypothesis and data, it is better to transform your model into the observed quantities rather than try to put the data into theoretical bins (star color and effective temperature are a characteristic pair). For complex situations, a Monte Carlo simulation is often (not always) the best way to do this - assume a model and calculate what the observers should see. There will, then, in effect be error bars on your theory as well as your observations, but this cannot be helped.

1.7 Conclusions

Statisticians and astronomers have been trespassing on each other's territories for as long as the territories have existed. In addition to the discovery of particular methods and concepts, we can find in this history several lessons. It is easier to analyze data you have taken yourself than other peoples (Mayer vs. Euler). Systematic errors generally exceed random ones (Maskelyne and many more recent examples). It is important to decide which hypothesis you are testing before you do the arithmetic, ideally even before you collect the data (archeo-astronomy and non-cosmological redshifts). And, finally, if as is nearly always the case, there is not a precise correspondence between the quantities you can measure and the ones in your hypothesis, it is best to transform theorists's units into observers' units, rather than the converse.

And the most important lesson is that the story is never completely told. Despite all these pages, I have not mentioned

- (1) Oscar Sheynin (1996 and many prior papers), who is the real expert on early astronomical statistics'
- (2) the early recognition of interstellar absorption by King (1914, working, as usual on "the" problem of stellar statistics);
- (3) all the good things that Simon Newcomb did (despite his role as Whiteman's "learned astronomer" and opposition to astrophysics), many of them statistical (corrections of coordinates for refraction, fluctuations of the solar cycle, recognition of the background light of the night sky as not being due to faint stars); or
- (4) Lambert (of the reflector), who despite Stigler's (1999) discussion of Bernoulli, Euler, and Old Max, arguably invented Maximum Likelihood (but did not use it for anything).

Acknowledgements: Special thanks to organizers Babu and Feigelson for finding something for me to do at the meeting and suggesting some references. Brenda Corbin of US Naval Observatory was enormously generous in finding and sending copies of papers not just older than UC Irvine but older than the present author.

1.8 REFERENCES

- [1] Abell, G.O. 1982. *Exploration of the Universe*, 4th Ed., Saunders, Holt Rinehart.
- [2] Badawy, 1968. *A History of Egyptian Architecture: The Empire*, Univ. of California.
- [3] Binney, J. and Tremaine, S.D., 1987. *Galactic Dynamics*, Princeton Univ. Press.
- [4] Bradley, J., 1728. *Phil. Trans. Roy. Soc.*, 34, 637-661.

- [5] Bradley, J., 1748. *Phil. Trans. Roy. Soc.*, 45, 1-45.
- [6] Brunt, D. 1917. *The Combination of Observations*, Cambridge Univ. Press.
- [7] Campbell, W.W. 1913. *Stellar Motions*, Yale Univ. Press.
- [8] Chandrasekhar, S. 1942. *Principles of Stellar Dynamics*, Chicago; Dover, 1960.
- [9] Eddington, A.S. 1914. *Stellar Movement and the Structure of the Universe*, London: McMillan.
- [10] Franklin, J. 2001. *The Science of Conjecture*, John Hopkins Univ. Press.
- [11] Giacconi, R. 1972. *Ann. NY Acad. Sci.*, 224, 149.
- [12] Hald, A. 1990. *A History of Probability and Statistics and their Applications Before 1750*, Wiley/Interscience.
- [13] Hald, A. 1998. *A History of Mathematical Statistics from 1750 to 1930*, Wiley.
- [14] Hampel, F. 1998. *Can. J. Stat.*, 26, 497.
- [15] Hoskin, M.A. 1963. *William Herschel and the Construction of the Heavens*, NY: Science History Publications.
- [16] Hoskin, M.A., Ed. 1999. *The Cambridge Concise History of Astronomy*, Cambridge Univ. Press.
- [17] Jeans, J. 1922. *Mon. Not. Royal Astro. Soc.*, 82, 122.
- [18] Jeffreys, H. 1930. *Theory of Probability*, Oxford: Clarendon Press.
- [19] Kapteyn, J.C. 1909. *Astrophys. J.*, 30, 284.
- [20] Kapteyn, J.C. 1922. *Astrophys. J.*, 55, 302.
- [21] King, E.S. 1914. *Harvard Obs. Annals*, 76, 1.
- [22] Krupp, E.C. 1988. In C. Ruggles, Ed., *Records in Stone*, Cambridge Univ. Press.
- [23] Kuiper, G.P. 1935. *Publ. Astro. Soc. Pacific*, 47, 38.
- [24] Malmquist, G. 1920. *Lund Medd.*, Ser. 2, No.22.
- [25] Malmquist, G. 1924. *Lund Medd.*, Ser. 2, No.32.
- [26] Maskelyne, N. 1762. *Phil. Trans. Roy. Soc.*, 52, 558.
- [27] McCuskey, S.W. 1965. In A. Blaauw & M. Schmidt, Eds., *Galactic Structure*, Univ. Chicago Press, 1.
- [28] Michell, J. 1767. *Phil. Trans. Roy. Soc.*, 57, 234-264.
- [29] Michell, J. 1787. *Phil. Trans. Roy. Soc.*, 74, 35-57.
- [30] Newcomb, S. 1886. *Amer. J. Math.*, 8, 343.
- [31] Newton, R. 1977. *The Crime of Claudius Ptolemy*, Baltimore.
- [32] Oort, J. 1932. *Bull. Astron. Neth.*, 6, 249.
- [33] Pearson, K. 1976. *The History of Statistics in the 17th and 18th Centuries*, Griffin & Co., Ed. E.S. Pearson.
- [34] Rao, C.R. 1997. In E. Feigelson & J. Babu *Statistical Challenges in Modern Astronomy II*, Springer Verlag.
- [35] Russell, H.N., Dugan, R.S. and Stewart, J.O. 1926. *Astronomy*, Boston: Ginn & Co.

- [36] Schaeffer, B.E. 2001. *J. Hist. Astron.*, 32, 1.
- [37] Scheuer, P.A.G.S. 1957. *Proc. Cam. Phil. Soc.*, 53, 764.
- [38] Schmidt, M. 2000. In P.C. van der Kruit & K. van Berkel, Eds. *The Legacy of J.C. Kapteyn*, Kluwer.
- [39] Scott, E.L. 1956. *Astron. J.*, 61, 190.
- [40] Sheynin, O.B. 1996. *The History of the Theory of Errors*, Egelsbach.
- [41] Stigler, S. 1986. *The History of Statistics*, Harvard Univ. Press.
- [42] Stigler, S. 1999. *Statistics on the Table*, Harvard Univ. Press.
- [43] Trimble, V. 1964. *Mitt. Inst. Orientforschung*, 10, 183.
- [44] Trimble, V. 1990. *Mon. Not. Royal Astro. Soc.*, 242, 79.
- [45] Trimble, V. 1992. *Scientometrics*, 20, 71.
- [46] Trimble, V. 1996. *Publ. Astro. Soc. Pacific*, 108, 1073.
- [47] Trumpler, R.J. 1929. *Lowell Observatory Bulletin*, 14, 154 (No. 420).
- [48] Trumpler, R.J. and Weaver, H.F. 1953. *Statistical Astronomy*.
- [49] van der Kruit, P.C. & K. van Berkel, Eds. 2000. *The Legacy of J.C. Kapteyn*, Kluwer.
- [50] van Rhijn, P.J. 1965. In A. Blaauw & M. Schmidt Eds. *Galactic Structure*, Univ. of Chicago Press, 27.