The Development of the Concept of the Best Mean of a Set of Measurements from Antiquity to the Present Day

Churchill Eisenhart [1]
66th ASA President
1971

1. Introduction

Presidents of the American Statistical Association in their presidential address commonly review the state of the Association or, at least, of some particular aspect(s) of the Association at the time of their presidencies, and then go on to predict future trends. I shall depart from this custom. I have chosen to talk not about the present and the future, but rather about how we got to the present from the past. In particular, I shall consider a central problem of statistical inference, the choice of the "best" mean of a set of measurements of a single quantity, and shall review practice and theory in this regard, from the dim past to the confused present. I am taking this opportunity to communicate to you some of the findings of research that I have been pursuing as a student of the history of statistics of measurement for a little over a decade.

In its simplest form, the problem of choosing the "best" mean is this:

> Given a number of "equally good" measurements of a single fixed quantity, what mean of their values should be taken as the best value of the unknown magnitude of this quantity afforded by these measurements?

A more general problem is:

> What mean should be taken when the measurements are NOT all "equally good"?

I shall be concerned principally with the simpler, more restricted, of these two questions; but will take up the broader problem from time to time.

There is an odd peculiarity of much historical writing and speaking that some of you may have noticed: The length of a historical account, or the definiteness of a historical statement, often tend to be related <u>inversely</u> to the amount of solid information available. Thus, I shall devote far more time and words to the early history of my topic than to more recent developments. This is justified in this instance, I believe, because many of you have participated in the creation or sharpening of the most recent developments, which are "well known" to you and others today; whereas, as I have discovered, the early history of my topic is not easy to come by and requires care in interpretation besides.

2. Preliminary Considerations

2.1 Some General Remarks -- Before we begin our historical journey together, let me set the stage, so to speak, by placing before you two quotations, the substance of which we shall do well to keep in mind as we proceed. The first is a statement by Professor Henry Guerlac of the

Department of History, Cornell University, on as irremediable shortcoming of man's historical record:

> To a greater extent than we often realize, what we can know about the past is what out ancestors—the participants in events or those who came soon after—determined that we should know. They placed in the intentional record—in annals, memoirs and commemorative inscriptions—those men and events which appeared to them as exceptional, striking and wholly outside the ordinary dull routine of private existence. In the main, they singled out for preservation in the collective memory those events which they saw to have markedly affected the way of life, the thoughts and actions, of the larger social groups and political entities: a tribe, a city-state, a nation or an empire. So it is that the main scaffolding and framework of our view of history consists of those deeds, thoughts and productions which others besides ourselves deemed worthy of preservation because of their effect upon man in society.
>
> -- Henry Guerlac [2]

The second, attributed to George Bernard Shaw by a speaker at the 50th National Conference on Weights and Measures,[3] elucidates a characteristic of ourselves that handicaps our interpretation of man's historical record:

> The fashion in which we think changes like the fashion of our clothes, and it is difficult, if not impossible, for most people to think otherwise than in the fashion of their own period.
>
> -- George Bernard Shaw

Efforts "to think otherwise than in the fashion" of our own time are made all the more difficult by changes in the meaning and usages of particular words. The etymologies of the terms "mean" and "average," which are especially relevant to our present historical journey, provide excellent illustrations. Today there is a tendency to use these terms more or less interchangeably, but their original meanings were very different.

Another obstacle to sure interpretation of European scientific and technical writings up to the 19th century is the lack of articles ("a" and "the") in Latin. This often makes it impossible to decide for sure whether a stated summary value is "<u>a</u> mean" or "<u>the</u> mean" of the corresponding set of measurements or observations, which are described but not given individually. If only "<u>a</u> mean," it may be simply some subjectively chosen value between the extremes of the set, or it may be some unspecified weighted mean of the individual observation; whereas, if "<u>the</u> mean," it is likely, early in our historical journey, to be the <u>midrange</u> (i.e., the "arithmetic mean" <u>between</u> the extremes), or, in more recent times, the <u>arithmetic mean</u> of all of the observations.

One more general comment. Some wag has remarked: "History is something that never happened, written by someone who was not there." This mischievous remark should, at least, remind us to beware of second- and third-hand accounts. In a study such as ours, we can, in effect, "be there" when we are able to examine original documents (or facsimile reproductions). But there remains, of course, the possibility of inaccuracy or incompleteness of the record; and of imputing a modern meaning or insight long before its time.

2.2 Means and Average [4] **–** The English economist and logician, William Stanley Jevons (1835-1882), complained a century ago that "Much confusion exists in the popular, or even the scientific, employment of the terms mean and average, and they are commonly taken as synonymous" ([5], p. 360). He went on to recommend ([5], pp. 362-363) that in scientific work the term mean be used when referring to an arithmetic mean of a set of measurements of a fixed quantity; and the word average, when referring to a "fictitious mean"[6] such as the arithmetic mean of the heights of the houses on a particular street, which serves to give an idea of their heights but may not be the height of any particular house. In both of these cases the expression "arithmetic mean" signifies, of course, the sum of the several individual measurements or heights involved divided by their number. In a few moments I shall point out that this is a comparatively recent extension of the original meaning of "arithmetic mean"…

3. The State of Affairs from Antiquity to the 17th Century A.D.

The taking of some definite mean – arithmetic mean, mode, median, midrange – of several measurements of a single fixed quantity to obtain a better value for its magnitude than was afforded by one or another of the individual measurements does not seem to have come a common practice until the 17th century A.D., and first appears in the latter half of the 16th, when a number of examples of the use of the arithmetic mean to this end can be found. Before the 17th century the picture is very fragmentary.

3.1 Absence of "Mean Taking": in Ancient Science and Technology – When I first began studying this matter, I expected to find a great many examples in antiquity. I had been brought up on the view that "astronomy [was] the most important force in the development of science since its origins sometimes around 500 B.C. to the days of Laplace, Lagrange, and Gauss" (Neugebauer [8]). So I fully expected that I would find some good examples of mean taking in ancient astronomy; and, perhaps, also in ancient physics. I have not found any. And I now believe that no such examples will be found in ancient science.

The reason is that quantitative science in antiquity was, to a large extent, mathematics and rested little on precise measurement. Great attention was devoted to mathematical details, and by the Greeks to mathematical elegance; but the observations involved were usually quite crude, thought very adroitly chosen… [9]

…the general practice in antiquity was to deduce a lot from very few data. For example, Professor Neugebauer has pointed out that tables "for the phenomena of Jupiter" from the late Babylonian period, say 240 to 40 B.C., which were computed several decades ahead, were nonetheless "based on a single observational element, the rest being derived therefrom in a strictly mathematical fashion."[11] "This," he adds, "conforms to a conscious tendency of the ancient to reduce the empirical data to the barest minimum, because they were well aware of the great insecurity of direct observations, especially for such major problems as the date of the first visible crescent or the reappearance of planets. All these phenomena are located near the horizon, where climatic and optical disturbances exercise a most pernicious influence" (Neugebauer [13]).

And in an earlier article on “Mathematical methods in ancient astronomy” he remarks: “In short, we can say that kinematics and spherical astronomy play a much greater role than empirical observations. The ancient astronomers were fully aware of the fact that the low accuracy of their instruments had to be supplemented by a mathematical theory of the greatest possible refinement. Observations are more qualitative than quantitative: ‘when angles are equal’ may be…

…decided fairly well on an instrument but not “‘how large’ are the angles,” says Ptolemy with respect to lunar and solar diameter (*Almagest V*, 14…).[14] Consequently, period relations over long intervals of time and lunar eclipses are the main foundations so far as empirical material is concerned; all the rest is mathematical theory… this holds for Greek as well as for Babylonian astronomy”(Neugebauer [16]).

The procedure just alluded to that Ptolemy (Claudius Ptolemaeus, fl . c . 150 AD) employed to obtain accurate values for the “apparent diameters” of the sun and moon provides a helpful illustration of the technique of supplementing crude observations with “mathematical theory of the greatest possible refinement.” Ptolemy says:

> But constructing ourself the four-cubit rod dioptra described by Hipparchus [fl. 161-126 BC], and making observations with it, we find the sun’s diameter everywhere contained by very nearly the same angle with no variation worthy of mention resulting from its distances. But we find the moon’s diameter contained by the same angles as the sun’s… only when during the full moons it is at its greatest distance from the earth… it was easy to see when each of the diameters subtends the same angle.. But how large they were seemed very doubtful to us… But once the moon at its greatest distance appeared to make an angle at the eye equal to the sun’s, by means of the lunar eclipses observed at that distance we calculated the angle subtended by the moon, and immediately we had that of the sun also.
>
> -- *Almagest V*, 14; English translation from {Ptolemy (151+) 1952}, pp. 171-172 [17]

In carrying out the step “by means of the lunar eclipses observed at that distances we calculated,” Ptolemy utilized Babylonian records of lunar eclipses “at that distance” that took place several centuries before his time <u>and</u> the whole intricate mathematical machinery of the then current theory of the moon’s motion; and reached the conclusion that

> The whole diameter of the moon subtends an arc of a great circle amounting to 0° 31 1/3’
>
> -- (<u>op</u>. <u>cit</u>., p. 173)

This agrees remarkably well with what I was taught in college, namely, that “the moon’s apparent diameter ranges from 33’30,” when nearest, to 29’21,” when most remote” (Russell, Dugan, and Stewart [18]).

Ptolemy’s *Almagest* is the main source for our knowledge of ancient astronomy. Ptolemy quotes observations of his own ranging from 127 to 151 A.D. and relies heavily on observations and methods of his predecessors, especially Hipparchus, who flourished almost three centuries before him, and whose works are lost, possibly partly the result of the fact that Ptolemy's great book superseded them and made them superfluous. (Neugebauer [16], p. 1013; Sarton [19]…)

…To summarize, it needs to be appreciated that observations and theory were differently related to each other in antiquity than they are today. Theory, the creation of the mind, reigned supreme, and observations were not understood as specifying "the facts" to which theory must conform, but rather as particular instances that were useful as illustrative examples in the explanation of theory, or as indicators serving as guideposts in the formulation of theory. (Compare Pannekoek [20]; and Sarton [21].)

To repeat, I have not found in ancient science any examples of the computation of a mean of two or more measurements of a single magnitude to obtain a more secure value; and I do not expect that any will be found. Nor have I found any in the records that have come down to us of day-to-day practical affairs.[22]

3.2 An Instance of Use of the Mode in Antiquity – About 20 years ago Professor W. Allen Wallis had the good fortune to discover an instance of the use of the <u>mode</u> of repeated counts in a measurement situation in the Peloponnesian war between Athens and Sparta (431-404 B.C.), as related by Thucydides (c. 460-c. 400 B.C.), one of the Athenian commanders:

> During the same winter [428 B.C.] the Plataeans, who were still being besieged by the Peloponnesians and the Boeotians, began to be distressed by failure of their supply of food, and since there was no hope of aid from Athens nor any other means of safety in sight, they and the Athenians who were besieged with them planned to leave the city and climb over the enemy's walls in the hope that they might be able to force a passage,… They made ladders equal in height to the enemy's wall, getting the measure by counting the layers of bricks at the point where the enemy's wall on the side facing Plataea happened not to have been whitewashed. Many counted the layers at the same time, and while some were sure to make a mistake, the majority were likely to hit the true count, especially since they counted time and again, and, besides were at no great distance, and the part of the wall they wished to see was easily visible. The measurement of the ladders, then, they got at in this way, reckoning the measure from the thickness of the bricks…
>
> -- Thucydides [25]

…During the remainder of 1963 and 1964, I corresponded or talked with a great many individuals [26] who might be aware of another (or, other) recorded instance (s) in antiquity of <u>multiple</u> or <u>repeated counting</u> followed by adoption if the mode. None could recall another instance, except in connection with <u>voting</u>, which, of course, involves <u>counting</u> (and, occasionally, <u>recounts</u> [27]) and adoption of the <u>mode</u>.

There is thus the nagging inference that this Peloponnesian War instance of "many" counting, and recounting, the layers of bricks, and adoption of the finding of the majority, i.e., the <u>mode</u>, may be an instance of the application in an unusual situation of the then prevailing rules of conduct in Athens and Sparta which "required that almost every important act be directed by a formal vote."[30] In this connection, Professor Tom Bard Jones of the Department of History, University of Minnesota, has drawn my attention to the accounts by Polybius (c. 205- 125 B.C.) and Livy (Titus Livius, 593 B.C.-17 A.D.) of the decisive tactics of the Romans in taking

Syracuse in 212 B.C., in which ladder-length for scaling the wall was determined from an estimate of its height based on “counting the courses”— but, while Livy’s account suggests that there may have been multiple or even repeated counting, there is absolutely no mention of adoption of the mode — quite the contrary!...

… I am not sure what we can conclude from all this beyond the following: there may well have been other instances in antiquity of multiple or repeated counting followed by taking the mode as the “true” value; but, in view of Professor Guerlac’s comments quoted earlier [2], such instances, if any, were not likely to be recorded for posterity unless some great good or evil ensued that made the even worthy of notice — in the above two instances, the “side” telling the story “succeeded”!

3.3 The Midrange as the Predecessor to the Arithmetic Mean – If there was a predecessor to the arithmetic mean as “best mean” – there certainly was no commonly employed predecessor – but if there was a predecessor at all, it must have been the <u>midrange</u>.

Al-Biruni (11th century A.D.) – Al-Biruni (Abu RayHan Muhammud ibn Ahmad al-Biruni, 973-1050+ A.D.), perhaps the greatest scientist of his day,[31] and one of the greatest of all time, refers to computation of the midrange of a set of measurements as if this were the customary “rule” in his day. In one instance he attributes this “rule’ to Ptolemy: --

> In the first treatise of the Almagest, Ptolemy stated that, for several years in succession, he observed [the sun’s zenith distance on the meridian at the times of the solstices]… At all times, he found it (the arc between the two solstices) to be forty-seven degrees, and more than two thirds, but less than three quarters of a degree. He assumed that this amounts approximately to what Eratosthenes had said, which was accepted by Hipparchus. <u>He said that because the rule is – for such a range with an upper limit and a lower one – to take the average amount between them.</u> Hence the amount (MS 78) given by Ptolemy is 47;42,30°; its half is 23;51,15°, but he constructed the tables of declination on the basis of 23;51,20°, in agreement with that assumed by Hipparchus and Eratosthenes, for if their third parts are rounded off, the declination comes to be this amount.
>
> -- al-Biruni [33] (Emphasis added.)

Ptolemy did not mention any such rule, or “give” the value quoted. All that Ptolemy said was--

> Now from such observations [of the sun’s zenith distance on the meridian] and especially from those made by us over several periods while the sun was near the tropics… we found the arc from the northernmost to the southernmost limit… to be always more than 47°40’ but less than 47°75’. And with this results nearly the same ration as that of Eratosthenes and as that which Hipparchus used. For the arc between the tropics turns out to be very nearly 11 of the meridian’s 83 parts.
>
> -- Ptolemy [34]

Thus, we see that al-Biruni "put words into the mouth" of Ptolemy that the latter never uttered.[35]

The midrange, or "average amount between" the limits that Ptolemy gives, is 47°42'30", as stated by al-Biruni. Also 11/83 of 360°, the value of the double obliquity (2e) that Ptolemy says was given by Eratosthenes (c. 273 – 192 B.C.) and adopted by Hipparchus (c. 162- c. 127 B.C), is equivalent to 47°42'39"2"' = 2(23°51'19"31"'); and Ptolemy actually used 23°51'20" (one-half of the foregoing rounded up) in the construction of his Table of Obliquity (e.g., [36]), entry in Table for 90°), thereby saying, in effect, that by his own observations he has confirmed this esteemed or traditional value. Britton has provided a fresh translation of the above passage, and has subjected Ptolemy's determination of the obliquity of the ecliptic to a detailed critique ([37]). Britton notes that Simon Newcomb's analysis (1895) of the decrease of the obliquity from antiquity to modern times gives 23°40'40" as the value of the obliquity (e) in Ptolemy day, so that the value Ptolemy used was too large by 21'20", and his error would have been reduced, but only slightly (by only 5"), had had used one-half of the midrange value for the double obliquity, namely, e = 23°51'15".

In Chapter V of his *Tahdid*…, al-Biruni again states the midrange rule and explains is purpose:

> As to the halving of the interval between the two times, it is a rule of procedure which has been adopted by calculators for the purpose of minimizing errors of observation so that the time calculated will be between the upper and lower bounds.
>
> --al-Biruni [38]

Assaying, to 1600 **–** Only single determinations are mentioned in the discussion of coin assaying in the so-called *Dialogus de Scaccario* ([39] Richard Fitzneale (?-1198), Treasurer of England and Bishop of London, of the procedures followed by the Exchequer in his day, in which he insists that the coin samples chosen for assay be well mixed so that they may "answer to the weight." Mention is made of duplicate and triplicate assays in the recently published 13th century English Mint Documents from the *Red Book of the Exchequer,* but no midrange or mean taking – rather "judgment should always be given for the assay which weighs the heaviest… because silver can easily be lost and can never be gained in the fire, so that judgment must be given where most silver is found" ([40], p. 82). No relevant details of assaying procedure are to be found in *De Moneta* [41], the treatise on money and coinage by the great French philosopher and scientist, Nicole Oresme (1323-1382). In the *Probierbuchlein* (1520+), "the first printed work on any aspect of metallurgy" ([42], p. 8), it is recommended that assays always be run in duplicate or triplicate "in order to be safe" ([42], para 45) – one assay may fail, or even two, but hopefully not all three (cf. Agricola [43]).

In *The Pyrotechnia of Vannoccio Biringuccio* (1480-1539), the wording [44], suggests that assays are to be made at least in duplicate. But nothing is said in either of these works about what is to be done if duplicate (or triplicate) results do not (all) agree. The implication is that, if the assays are successfully carried out, the resulting "beads" of pure silver will weigh exactly the same. This is explicitly stated in *De Re Metallica* (1556) of Georgius Agricola (Georg Bauer, 1494-1555): "If neither [bead] depresses the pan of the balance in which it is placed, but their weight is equal, the assay has been free from error; but if one bead depresses its pan, then there is an error, for which reason the assay must be repeated" ([43], p. 252).

Astronomy, to 1600 – To date I have not examined any European works on astronomy prior to the great *De Revolutionibus Orbium Coelestium*_ [45] by Nicolaus Copernicus (1473-1543), Canon of the Cathedral at Frauenberg in East Prussia (now Poland), which changed the course of astronomy. I have looked through the "Great Books" English translation {c 1952} several times without finding any instance or mention of taking a "mean" of any kind (midrange, arithmetic, mean, etc.) of two or more determinations of a single quantity. At one point in [45] (Book VI, Chapter 7 ({c 1952}, p. 287) Copernicus does recommend using the midvalue between the extremes of a variable astronomical quantity, to minimize computational labor and simplify exposition, "where, indeed, the extremes would not have made any manifest difference" ("ubienim extrema non fecerint apertam differentiam" [45], p. 430); but this, of course, is not the same as taking the midrange of multiple determinations of a single fixed quantity in order to obtain a more secure value.[46]

Before leaving Copernicus, let me note that he also wrote a treatise on coinage [48], at the invitation of the King of Poland to serve as a basis for currency reform in the Prussian states of Poland. I have examined a French translation [c 1864] and found nothing on assaying procedure and no mention of "mean taking" of any kind.

To my great disappointment, I have not, to date, found in astronomical writings any clear cut (or even probable) examples of the use of the midrange of more than two measurements of a single quantity as a more secure value, from al-Biruni's day up to 1650 A.D., say, by which time the arithmetic mean was unquestionably in use.[49] But I have found some later (at least probable) examples…

…***Navigation to 1600*** – Practically everything that is known about this history of navigation among Mediterranean and Western European peoples from antiquity to the close of the 18th century is contained in *The Haven-Finding Art* by the late E.G.R. Taylor,[51] "a history of navigation 'from Odysseus to Captain Cook,' which in both scope and authority is perhaps the most definitive book of its kind." A useful adjunct is *The Art of Navigation in England in Elizabethan and Early Stuart Times* by D.W. Waters,[52] by virtue of its precise notes and references and detailed appendices.[53] From a study of the early chapters of these volumes it is clear that navigational practices on the part of Mediterranean and Western European seamen were completely non-numerical up to only a few years before Christopher Columbus's first voyages to America. Consequently I consider it practically certain that no "mean taking" of any kind was used by these seamen before the close of the 15th century.

It was the far-flung ocean voyages of the Portuguese in the 15th century that focused attention on the need for dependable methods for determining a ship's position at sea.[55] The first step toward mathematization of navigation was astronomical but non-numerical. As the sea captains of Portugal's Prince Henry the Navigator (1394-1460) sailed south along the West Coast of Africa, they saw the North Star drop lower and lower in the northern sky behind them and used its altitude as an indication of their distance south of Lisbon. Prince Henry's learned advisors formalized this procedure and introduced the new navigational technique of instrumental observation of the altitude ("altura") of the North Star.[57] "But… at first, as pilots did not know how to use a scale of 'degrees,' and had never thought in terms of 'latitude,' quadrants were marked with important place names against particular parts of the scale, and the pilot could thus recognize his position by the fall of the plumb line alone." ([12], p. 159).[58]

It was the dropping out of sight of the North Star as the Portuguese seafarers approached and crossed the equator, and continued their explorations southward, that forced the shift from place names on quadrants to latitudes in degrees, and thus began the arithmetization of navigation. The requisite *Rules, or Regiment, of the Sun* – for translating observed altitudes of the noonday sun (in degrees) into (degrees of) latitude north or south of the equator – were prepared about 1485 by a commission of astronomers and mathematicians appointed by Prince Henry's grandnephew, King John II of Portugal. The commission also simplified the astronomer's or astrologer's planispheric astrolabe, "by leaving out the parts not absolutely needed and produced a plain astrolabe of wood and iron" ([60]) for measuring the altitude of the sun or North Star, and provided a set of corrections between +3.5° and -3.5° to permit more accurate translation of observed altitudes of the North Star in degrees of latitude north of the equator. A copy of these *Rules* for converting altitudes of the sun or North Star into degrees of latitude was apparently used by Christopher Columbus on his voyages to America (1492-1504) and by Vasco de Gama on his voyage to India around South Africa's Cape of Good Hope in 1497-1498.

The rolling and pitching motions of a ship made accurate measurements of the altitude of the sun or a star with a quadrant or astrolabe difficult, and often impossible, on shipboard.[61] Consequently, seamen went ashore to take altitude observations, if possible. But when a need for knowledge of latitude arose at sea, especially when the safety of the ship and cargo and every man aboard was in jeopardy, they simply had to do as best as they could. Therefore, I had hoped to come upon some 16th century instances of "mean taking" at sea in order to make the best of a set of discordant observations. But, to my great disappointment, I have not, to date, found any clear-cut examples of "mean taking" of any kind in writings on navigation until the last decade of the 16th century, where I have found two…

…***a. 1595 (Harriot)*** – Thomas Harriot [64] suggested in his "Instructions for Raleigh's Voyage to Guiana, 1595," which date from late 1594 or early 1595, that when taking the altitude of the sun on shipboard with an astrolabe or sea ring,[66] the midpoint between the extreme readings, i.e., the midrange, be taken as the true reading…

…A few special features of the midpoints or midranges involved in this recommendation should be noticed. The sea-ring procedure calls for reading "the middle of the play" directly; and the first of the two astrolabe procedure calls for adjusting the Index (i.e., the alidade) until the oscillation of the light spot (from the light through the upper sight) is centered on the hole of the lower sight, and then reading the location of the Index points on the scale. In both of these cases, the midpoint or midrange is "observed" directly, without computation from the values of the "extremes." In contrast, the second astrolabe procedure involved finding and reading the "extremes" and then computing the midpoint or midrange. Finally, in all three cases, the "sample size." i.e., the total number of "individual" momentary positions of the light spot or index at least implicitly involved, is indefinite.

… ***b. 1599 (Wright)*** – Inasmuch as Harriot's contributions to navigation, developed "for the exclusive use of Raleigh's navigators" ([67], p. 590), were never published, I was pleased to find a somewhat similar recommendation in *Certaine Errors in Navigation…Detected and Corrected* (1599) by Edward Wright,[68, 69] "the most influential and oft-quotes treatise on nautical practice of the era" ([71], p. 45). The pronouncement of interest to us occurs in connection with Wright's discussion of measurement of the variation of the compass as a navigational tool.[72] If, as he believed, compass variation could be used in combination with latitude determinations to indicate a ship's position at sea, by comparing observed with tabulated values, it was clearly important that determination of compass variation on shipboard be carried out as accurately as possible. Therefore, in order "that others that shall go about hereafter to observe the variation (at sea especially) may bee the more circumspect to foresee and preuent all causes of error herein," he advised:

> Exact truth amongst the inconstant waues of the sea is not to bee looked for, though good instruments bee neuer so well applied. Yet with heedfull diligence we may come so neare the trueth as the nature of the sea, our sight and instruments will suffer us.…
>
> -- Edward Wright [68]
> verso of lN1

…Unfortunately, Wright does not give any numerical examples of the applications of this dictum in either his original (1599) or expanded (1610) edition [73]. Consequently, it is not possible to be absolutely certain that he is recommending use of the median, as he appears to be. However, his admonition that "neither are [the observation] all by & by to be rejected" seems to me to support the inference that he is recommending the "middlemost: observation, i.e., the median, and not the "middle place," i.e., the midrange: correct identifications, of sufficient central observations to be on the safe side as more observations are taken and the number of outliers discarded in each direction; whereas computation if the midrange requires knowledge of only the largest and smallest observations throughout the period of observation…

4. The Rise and Fall of the Principle of the Arithmetic Mean

By far the best know statement of this Principle is that which Gauss gave as the starting point of his derivation of the (Normal) Law of Error in article 177 of his great astronomical treatise, *Theoria Motus Corporum Coelestium…* ([74]). What he said, in English translation, was this:

> It has been customary certainly to regard as an axiom the hypothesis that if any quantity has been determined by several direct observations made under the same circumstances and with equal care, the arithmetical mean of the observed values affords the most probable value, if not rigorously, yet very nearly at least, so that it is always most safe to adhere to it.[75]
>
> -- Gauss [74], p. 258

Gauss clearly makes no claim to having originated this "hypothesis" that he adopts as an "axiom." Quite the contrary; he regards it as traditional. And rightly so, because, as we shall see, the practice of adopting the arithmetic mean of a number of "equally good" measurement of some quantity as the "best" value of this quantity afforded by these particular measurements certainly predicts Gauss's birth (April 30, 1777) certainly by more than one, and, perhaps, by as much as two centuries.

4.1 From No Mention of "Mean Taking" of Any Kind to Explicit "Taking of the Arithmeticall Meane" "for the True Variation" in the 16th-17th Century Writings on Magnetic Declination – The detailed discussion by Gustav Hellmann [76] of early geomagnetic observations ([77]), together with his facsimile reproductions of practically all of the important writings on geomagnetic phenomena up to 1635 (Hellmann [78], Gellibrand [79]), have enabled me to trace the evolution of "mean taking" in writing on magnetic declination (or variation of the compass), from no mention of "mean taking" of any kind early in the 16th century, through seeming adoption (without explicit mention) of the arithmetic mean toward the close of the 16th century, to explicit mention of "taking the Arithmeticall meane" in the mid- and late 17th century. To date I have not succeeded in unfolding the evolution of "mean taking" so neatly in any other area of science...

….the discovery of marked differences in the observed direction and magnitude of the variation of the compass (or magnetic declination) in different parts of the world toward the close of 15th century led to the hope early in the 16th century that the isogonic (Greek *isogonios*: "having or pertaining to equal angles") lines or equal magnetic declination over the surface of the earth would exhibit a stable and orderly pattern which, in conjunction with the parallels of latitude, would enable a ship to determine its precise position at sea from observation of declination and latitude alone, by-passing the need for direct determination of longitude. Therefore, measurement of magnetic declination was pursued with diligence throughout the 16th and the early part of the 17th centuries.

Many of the early determinations of magnetic declination obtained merely by noting the deviation of the magnetic needle from "north" as determined simply by sighting the North Star (Polaris) over the compass dial, were subject to great uncertainty, especially in the case of observations made on shipboard. "That in this way no great accuracy could be attained is self-evident. It is also to be questioned whether the movement of Polaris… about the North Pole [in] a circle of about 5 degree in diameter… was always taken into account." ([77], p. 81) Indeed, declination values obtained by different pilots at the same place not only agreed poorly with each other in magnitude, but often contradicted each other with respect to direction, so "that doubts of the correctness of the magnetic declination arose everywhere…" (loc. cit.). Clearly a prerequisite to accurate measurement of the deviation of a compass needle from the north at any particular place is accurate knowledge of the direction of "north" at that place; and this is what was lacking, especially on shipboard.

The customary procedure for finding the north-south direction or meridian line at a particular place before the advent of the "north-pointing" magnetized needle was to bisect the angle between morning and afternoon shadows of equal length cast be a gnomon.[80] In 1525, Felipe Guillen, an apothecary of Seville, presented to King John III of Portugal a portable instrument ("brújula de variacón", i.e., "sea compass for variation") that he had devised for measuring the magnetic declination at a particular place on land or at sea. It consisted of a small sun-dial setup with a magnetic needle and an azimuth scale graduated from 0° to 180° clockwise (from N. through E. to S.) and counterclockwise (from N. through W. to S.). ..

…Searching through early issues of the *Philosophical Transaction*s [84] of the Royal Society, I came upon the following clear-cut example of "taking" the arithmetic mean as the "best" approximation to the truth.

1668 (D.B.) – The issues of the *Philosophical Transaction*, dates "Monday, July 13, 1668," contains "An Extract of a Letter, written by D.B. to the publisher" [85] presenting a table of five magnetic declination measurements made near Bristol on June 13, 1666, by Capt. Samuel Sturmy,[86] "an experienced Seaman, and a Commander of a Merchant Ship for many years," who took them "in the presence of Mr. Staynred,[87] an ancient Mathematician, and others."

> In this table, he [Capt. Sturmy] notes the greatest… difference to be 14 minutes; and so taking the mean for the true variation, he concludes it then and there to be just 1 degree 27. minutes.
>
> -- D.B [85], p. 726

The five individual determinations "in this table" are:

1 degree 22', 1 degree 36', 1 degree 34', 1 degree 24', 1 degree 23'.

The exact arithmetic mean of these five values is 1 degree 27 4/5'; the median is 1 degree 24'; the midrange is 1 degree 29'. The "mean" taken "for the true variation" was the arithmetic mean indubitably!

5. References

[[1] Because the text of Churchill Eisenhart's presidential address was not available, this version of the talk is a synopsis of a fuller paper on the topic. The speech was presented at the 131st Annual Meeting of the American Statistical Association at Colorado State University, Ft Collins, CO, on August 24, 1971. Editing of the text was provided by Fritz Scheuren and Wendy Alvey.**]**

[2] Guerlac, Henry (1963), 798.

[3] Smith, M.D. (1966), p. 145. Thus far I have not succeeded in finding the precise location of this statement in Shaw's writings or in verifying that Shaw is its author.

[4] This subsection should be skimmed or perhaps skipped on first reading. It brings together in one place various facts about means and averages mentioned only briefly or merely alluded to at various points in the oral presentation together with additional relevant information needed to complete the story.

[5] Jevons, William Stanley.

[6] I am giving here the example used by Adolphe Quetelet (1796-1874) to make the same distinction (Quetelet [7]), because it is much simpler than Jevon's example involving the mean

density of the earth. Quetelet elected to reserve the word mean for the first case and employ the full expression, arithmetic mean, in the second.

[7] Quetelet, Adolphe (1849), p.42.

[8] Neugebauer, Otto (1952), p. 2.

[9] This important fact had somehow escaped my attention in my earlier reading. Professors Aaboe and Price at Yale have brought it out forcefully in a fascinating article [10]. They demonstrate how amazingly accurate numerical values for astronomical parameters could be derived from single crude but crucial observations corresponding to special or extreme circumstances that were especially favorable or decisive and point out that much ingenuity was employed in the choice of just what to observe in the application of mathematical machinery and in the solution of the basically mathematical problems involved.

[10] Aaboe, Asger and Price, Derek J. de Solla (1964), The Derivation of Accurate Parameters from Crude but Crucial Observations.

[11] Such practices were not restricted to antiquity. The late Professor Eva G.R. Taylor (1880-1966) mentions a tide table dating from early in the thirteen century AD and associated with Matthew Paris (c 1200-1259 AD), which states that, "at new Moon high tide is said to be at 3 hours 48 minutes, on the second day at 4 hours 36 minutes, on the third at t hours 35 minutes, and so on." [12] Miss Taylor comments, "It is, in fact, mechanically built up from a single observation according to the rule accepted by astronomers that the daily retardation was 48 minutes. It is a scholar's, not a sailor's, table. When John Flamsteed drew up his tide table for London Bridge in 1676, he found the figure highly variable, the retardation sometimes under 30 minutes, sometimes an hour or more… Theory rather than observation was still the rule in the learned world of Matthew Paris's day."

[12] Taylor, Eva G.R. (1956), Flod at London, pp. 136-137.

[13] Neugebauer, Otto (1954), p. 801.

[14] The original Greek title translates as *Mathematical Syntaxis* or *Mathematical Collection*. To distinguish it from lesser astronomical treatises of other authors, later commentators dubbed it *The Great Collection*, which passed into Arabic as *al Magisti*, "the Greatest." In time this became *Almagest*, by which name it has been known ever since. [15]

[15] Heath, Thomas L. (1931) *A History of Greek Mathematics,* reprinted in 1963, pp. 402-403.

[16] Neugebauer, Otto (1948), p. 1015.

[17] Ptolemy, (151+) *Almagest V*, English translation (1952), pp. 171-172.

[18] Russell, H.N.; Dugan, R.S.; and Stewart, J.Q. (1926), *Astronomy,* New York: Ginn & Co., p.163.

[19] Sarton, George (1954), *Ancient Science and Modern Civilization,* New York: Harper & Bros., pp. 41-42.

[20] Pannekoek, Anton (1961), *A History of Astronomy*, New York: Interscience Publishers, p.150.

[21] Sarton, George (1959), *A History of Science. Hellenistic Science and Culture in the Last Three Centuries*, Cambridge: Harvard University, p. 57.

[22] For a while I believed that one had been found. .Julian Lowell Coolidge (1873-1954) [23] refers, without precise identification, to a "set" of problems in a 1935 paper of Francois Thureau-Dangin (1872-1944) and comments: "Here a field is divided into rectangles, right triangles, and rectangular trapezoids, as described above. The total area is calculated from two different sets of measures and the divergent results are averaged. It is interesting to note that the Babylonians realized that there must be slight errors of observation and sought means to obviate them." To date, however, I (and a number of others much more expert in such matters than I) have been unable to find any such case as he described. It is certainly not in the paper he cited [24]…

[23] Coolidge, Julian Lowell (1940), *History of Geometrical Methods,* Oxford: Oxford Press, p. 5.

[24] Thureau-Dangin, Francois (1935), *La Mesure des Volumes D'après une Tablette Inedite du British Museum*, Revue d'Assyrologie 32.

[25] Thucydides (431 B.C.E.), *History of the Peloponnesian War,* Book III, para. 20, Greek text , p. 30, 32; English translation by Charles Forster Smith (1919), London: London Heinemann, p. 31. 33.

[26] See full paper for list of contacts.

[27] Professor A.H.M. Jones commented that inasmuch as "the Athenians voted by a show of hands, …there was not exactly a recount, as people could change their votes;" [28] and he cited an instance of a second vote taken in the Athenian assembly in 406 BC, described by Xeonophon (c 430-c 350 BC) in [29], in which the second vote, which reversed the decision, was taken after an objection was made to the legality of the motion adopted in the first vote. Perhaps some reader will be able to send me an earlier and "cleaner" example.

[28] Jones, A.H.M. (1958), *Athenian Democracy,* New York: Frederick Praeger, Inc.

[29] Xenophon, *Hellenica*, Book I, Chapter VII, para 34.

[30] "In ancient Greece and Italy, the institution of suffrage already existed in a rudimentary form at the outset of the historical period. In the primitive monarchies, it was customary for the King to invite pronouncements of his folk on matters in which it was prudent to secure its assent beforehand. In these assemblies, the people recorded their opinion by clamouring (a method

which survived in Sparta as late as the 4th century BC), or by the clashing of spears on shields… the word “suffragium” meaning literally a responsive crash… in the days of their full political development, the communities of these countries (e.g., Athens in the Age of Pericles, 443-429 BC) had firmly established the principle of government according to the will of majorities and their constitutions required almost every important act to be directed by a formal vote.” -- *Encyclopedia Brittanica,* Vote and Voting, 11th edition, vol. 28, p. 216.

[31] Al-Biruni (“The Master”) wrote important treatises on astronomy, geodesy, mathematics, mechanics, mineralogy, and pharmacology, among others. For a full and interesting biography, see [32].

[32] Kennedy, E.S. (1970), *Dictionary of Scientific Biography*, New York.

[33] Al Biruni (1018 A.D.), *al-Amakin,* Chapter III, translation by Jamil Ali, pp. 59-60.

[34] Ptolemy, *Almagest I,* 12; Taliaferro translation, p. 26.

[35] Professor Neugebauer told me in the spring of 1966 that he knew of no edition of the *Almagest* that contains the rule statement that al-Biruni attributes to him, but that it is possible that there once existed, and may still exist somewhere, an Arabic version of the *Almagest* containing such a statement, because the Arabic scribes and scientists were prone to “improving” manuscripts they copied. Professor Neugebauer was quite positive that there was no such “rule” in Ptolemy’s day…

[36] Ptolemy, *Almagest I,* 15, p. 31,

[37] Britton, John Phillip ( ), Chapter 1, pp. 1-2.

[38] Al-Biruni (before 1025 A.D.), *Tahdid…*, Chapter V, translation, p. 168.

[39] Fitzneale, Richard (1902), *Dialogus de Scaccario,* eds. A. Hughes, C. G. Crump, and C. Johnson, pp. 36-43.

[40] Oresme, Nicole (1838), *Red Book of the Exchequer,* translation by Joseph Hunter, London: Pickering.

[41] Oresme, Nicole (1956), *De Moneta de Nichole Oresme,* translation by Charles Johnson, London: Nelson.

[42] (1520+), *Probierbuchlein,* p. 8.

[43] Agricola, Georgius (1556), *De Re Metallica*, translated by Herbert Clark Hoover and Lou Henry Hoover (1912), London: Salisbury House, p. 241.

[44] Biringuccio, Vannoccio (1540), *The Pyrotechnia of Vannoccio Biringuccio,* translated by C.S. Smith and Martha Teach Gnudi (1942), Dover: Phoenix, p. 139.

[45] Copernicus, Nicolaus (1543), *De Revolutionibus Orbium Coelestium* , Basel: Officina Henricpetrina; translated to *On the Revolutions of the Celestial Spheres* (1952).

[46] In his calculation of the angle of obliquation of Venus (loc. cit.), Copernicus adopts Ptolemy's values [47] for Venus's distances at apogee and perigee, but in transformed units, so that "the greater distance, at apogee, is 10,208" and the lesser, as perigee, is 9,792" "and also the mean distance is 10,000"…

In the "Great Books" English translation… (c 1952]), p. 827:

> "…which Ptolemy decided to assume in this demonstration, as he wished to avoid labour and difficulty and to make an epitome, for where the extremes do not cause any great difference, it is better to use the mean."

(While this is not a strictly accurate translation, it is in no way misleading.)…There is one more point of interest here: while Ptolemy uses the middle value (60^p in his units) in his "demonstration," he does so without comment. The explanation is all Copernicus's own addition.

[47] Ptolemy, *Almagest XIII*, 4; p. 449.

[48] Copernicus, Nicolaus (1526), *Monete Cudende Ratio,* first printed in 1816; translated into French c 1864.

[49] I am deferring to the next section discussion of the "mean taking" practices of Tycho Brahe (1546-1601), "the prince of astronomical observers" ([50], p. 270), because some of his data and associated summary values that I examined fall among the ambiguous cases that I have come upon in the period 1550-1650 – "ambiguous" because the arithmetic mean is NOT mentioned explicitly and the summary value given, in consequence of being recorded more coarsely than the individual observations, could be a rounded value of more than one of the simple "averages" (arithmetic mean, median, midrange, or mode).

[50] More, ( ) on Tycho Brahe, p. 270.

[51] Taylor, Eva Germaine Rimington (1956), *The Haven-Finding Art: A History of Navigation from Odysseus to Captain Cook,* London: Hollis & Carter.

[52] Waters, David Watkin (1958), *The Art of Navigation in England in Elizabethan and Early Stuart Times,* London: Hollis & Carter.

[53] A concise yet remarkably complete chronological summary of the principal elements of the Three Ages of Pilotry ("primitive," "quantitative," and "mathematical" navigation) in "the western part of the Old World" has been provided by Joseph Needham in [54].

[54] Needham, Joseph (1971), Three Ages of Pilotry, in *Science and Civilization in China*, London: Cambridge University Press, volume 4, Part III, pp. 554-560.

[55] The exact location of a ship at sea is uniquely determined by its latitude and longitude… the Portuguese pioneers of the 15th century solved the “latitude problem,” but a general solution to the technically much more difficult “longitude problem” was not found until the latter half of the 18th century – see, e.g., [56].

[56] Brown, L.A. (1956), The Longitude, in *The World of Mathematics*, J.R. Newman (ed.), vol. 2, New York: Simon and Shuster.

[57] The latitude… of a place on earth is equal to the altitude of the northern, or southern, celestial pole as seen by an observer at the place. Inasmuch as the North Star … revolved around the northern celestial pole in a circle of about 3.5 degrees in the late 15th century, the latitude of a ship north of the equator was indicated to a very good approximation by the altitude of the North Star.

[58] For a sketch of a marine quadrant of 1492, and a vivid description of the manner and difficulties of its use at sea, see [59], p. 185.

[59] Morison, Samuel Eliot (1942), *Admiral of the Ocean Sea: A Life of Christopher Columbus*.

[60] Prestage, Edgar (1933), *The Portuguese Pioneers*, London: A&C Black, p. 319.

[61] Thus, Columbus, homeward bound on his first voyage to America, noted in his journal entry for 3 February 1493:

> “The North Star appeared very high, as on Cape St. Vincent; but couldn’t take the altitude with the astrolabe or quadrant, because the rolling wouldn’t permit it.” ([62])…

The difficulties and errors of altitude measurement on shipboard were not entirely overcome until the advent of the reflecting quadrant (later octant and sextant), invented independently in 1730 by John Hadley (1682-1744) in England and Thomas Godfrey (1704-1749) in Philadelphia. (See, e.g., [63] or [12], pp. 256-258…)

[62] Morison, Samuel Eliot (1963), *Journals and Other Documents on the Life and Voyages of Christopher Columbus*, New York: Heritage, p. 160.

[63] Cotter, Charles H. (1968), *A History of Nautical Astronomy*, London: Hollis & Carter, pp. 77-81.

[64] Thomas Har(r)iot (1560-1621), “one of the founders of algebra as we know the science today” ([65], p. 388) entered the service of Sir Walter Raleigh in 1579 as mathematical and scientific advisor, especially to solve mathematical problems arising in navigation and to correct errors in current navigational practice….

[65] Smith, David Eugene (1923), *General Survey of the History of Elementary Mathematics*, New York: Dover Publications.

[66] See, e.g., [63], fig.2, p. 63.

[67] Waters, David Watkin (1599), *The Art of Navigation in England in Elizabethan and Early Stuart Times*, New Haven, p. 590.

[68] Wright, Edward (1599), *Certaine Errors in Navigation…,* London: Valentine.

[69] Edward Wright (1558-1615) "was a Fellow (1587-1596) of Caius Collage, Cambridge, who was induced by the Earl of Cumberland about 1589 to apply his mathematical studies, particularly to navigation…. He designed, described, and made mathematical instruments of many types [and he] gave William Gilbert (1540-1603) considerable assistance in his great work *De Magnete* [70]…

[70] Gilbert, William (1600), *De Magnete*, London: Chiswick Press.

[71] Taylor, Eva Germaine Rimington (1954), *The Mathematical Practitioners of Tudor and Stuart England*, Cambridge: University Press, p. 45.

[72] During the 15th century, it became recognized that a magnetic needle does not, in general, point directly to the north, but deviates from the true meridian by a small angle that varies from place to place. This variation of the compass (…or magnetic declination, as it is now termed) was easterly over Western Europe. Then, "the voyages to the Americas and to India… brought …[the news] that after a ship had passed [west of] the Azores or, alternatively, had rounded the south of Africa [to the east], there was a change from 'north-easting' to an increasing 'north-westing'" ([51], p. 173)…measurement of compass variation provided a means of solving or by-passing the longitude problem…and would enable a ship to determine its position at sea…

[73] Wright, Edward (1610), *Certaine Errors in Navigation…,* 2nd ed., London: Valentine.

[74] Gauss, Karl Friedrich (1809), *Theoria Motus Corporum Coelestium…*, Hamburg: Fred. Perthes & I.H. Besser.

[75] …The English translation ([74] 1857, p. 258) is by Charles Henry Davis (1807-1877), first Superintendent of the American Ephemeris and Nautical Almanac Office (1849-1856) and twice Superintendent of the U.S. Naval Observatory (1865-1867, 1874-1877).

[76] (Johann Georg) Gustav Hellmann (1854-1939), Professor of Meteorology (from 1886) and Direct (1907-1922) of the Meteorology Institute, University of Berlin.

[77] Hellmann, Gustav (1899), Wetterprognosen und Wetterberichte des XV und XVI Jahrhunderts, in *Neudrucke von Schriften und Karten über Meteorologie und Erdmagnetismus*, (ed.), Berlin: A. Asher & Co.

[78] Hellmann, Gustav (1898), Ueber Luftelektricität 1746-1753, *Neudrucke von Schriften und Karten über Meteorologie und Erdmagnetismus*, (ed.), Berlin: A. Asher & Co.

[79] Gellibrand, H. (1635), *A Discourse Mathematical on the Variation of the Magnetic Needle Together with its Admirable Diminution Lately Discovered*, London: William Jones.

[80] It is described clearly, for example, by the Roman architect Vitruvius (c 88 – c 26 BC) in his *De Architectura Libri Decem* … [81] Commonly known as the Indian circles method, this procedure has been traced back to an Indian astronomical work dating from about 400 BC ([82]). An older and more accurate method, but requiring a wide and absolutely level horizon and accurate sighting, is based on analogous bisection of the angle between the points of rising and setting of some particular star and was used by the ancient Egyptians to fix the orientation of the temples and pyramids ([Edwards], pp. 258-260).

[81] Vitruvius (c 1486), *De Architectura Libri Decem*, 1st printed in Rome; translated to English by M. H. Morgan (1914).

[82] Kiely, Edmund (1947), *Surveying Instruments: Their History*, New York: Columbia University, pp. 37, 61-62, 281-282.

[83] Edwards, I.E.S. (1987), *The Pyramids of Egypt*, London: Pelican Books.

[84] Publication of which began in 1665, as first partly as a private profit-seeking venture of the Society's Secretary, Henry Oldenberg (c 1615-1677).

[85] D. B. (1668), An Extract of a Letter Written by D.B. to the Publisher, Concerning the Present Declination of the Magnetick Needle, and the Tydes, *Philosophical Transactions*, May 23, 1668, 3, 726-727.

[86] Captain Samuel Sturmy (1633-1669). For further details, see [71], Biog. 265a, Work 329.

[87] Philip Staynred (Strandridge), (fl. c 1621-1669). For further details, see [71], Biog. 145 and Work 328.