The Depth of Association Between the Ancient Star Catalogues

Dennis W Duke

Florida State University

The primary sources of star data from antiquity are Hipparchus' *Commentary to Aratus*[1] and Ptolemy's Almagest[2]. Both sets of data have errors, and it has long been of interest to try and establish associations between the data sets by looking carefully at correlations between the two sets of errors[3].

A careful analysis[4] of the statistical and systematic errors in the rising and setting phenomena shows that the degree of association is very likely substantial. In particular, if we concentrate on the correlations between the statistical errors in the rising and setting phenomena, we see that the size of the correlations is easily understood by means of a simple model: the Almagest errors are ε_i, where ε has mean zero and variance σ_A^2, while the Commentary errors are $\varepsilon_i + \eta_i$, and these have mean zero and variance σ_C^2, while ε and η are completely uncorrelated. The Commentary and Almagest errors share the ε errors, but the added η errors account for the empirical fact that the variance σ_C^2 in the Commentary errors is larger than the variance σ_A^2 in the Almagest errors.

The idea behind this model is that there was an Hipparchan catalog of star positions which is now lost, but which was available to Ptolemy for copying. The data that appear in the Commentary were also related to this Hipparchan data, but were derived from it. The most plausible source of the added η errors is, of course, that Hipparchus used a globe to obtain his numbers for the phenomena, but the source is not important to the present argument. Thus, the model implicitly assumes that the errors on the values in the Commentary will reflect whatever errors were in Hipparchus' lost catalog, and then in addition whatever errors result from his calculation of the phenomena using those lost coordinates.

Under these assumptions, the correlation r between the two sets of errors is just $r = \sigma_A / \sigma_C$, which can have any value between zero and unity, even when, by the model assumption, every star has been copied (although, of course, not from the Commentary, but from a presumed earlier source). Since all three of the numbers r, σ_A, and σ_C are measured from the data, the model has no adjustable parameters and is easy to check. The results are given in Table 1. We see that overall the agreement with the data is within reasonable expectations for a zero-parameter model, and supports the idea that the Commentary data and the Almagest data share a common origin[5]. Note also that this model provides an explicit counterexample to the assertions by Swerdlow[6] and Evans[7] that r is directly related to the percentage of copied coordinates.

James Evans suggested[8] to me a simple extension of the model, and it turns out that the extended model allows an estimate of the fraction of stars copied by Ptolemy. The extended model is specified by assuming the following error relationships:

a lost Hipparchan catalog	ε_i	$i = 1..N$
the Commentary	$\varepsilon_i + \eta_i$	$i = 1..N$
the Almagest catalog	ε_i	$i = 1..n$
	ζ_i	$i = n+1..N$.

Thus for the Almagest catalog we assume that Ptolemy copied n of the N total stars, but measured afresh, with independent errors, the remaining N-n stars. These errors all have zero mean and are uncorrelated, so $\langle\varepsilon\rangle = \langle\eta\rangle = \langle\zeta\rangle = \langle\varepsilon\zeta\rangle = \langle\varepsilon\eta\rangle = \langle\eta\zeta\rangle = 0$, and the fraction f of stars copied is n/N.

Then it is easy to show that the variances and the single non-zero correlation r between the Commentary and ASC errors are

$$\sigma_H^2 = \langle \varepsilon^2 \rangle$$
$$\sigma_C^2 = \langle \varepsilon^2 \rangle + \langle \eta^2 \rangle$$
$$\sigma_A^2 = f\langle \varepsilon^2 \rangle + (1-f)\langle \zeta^2 \rangle$$
$$\sigma_A \sigma_C r = f\langle \varepsilon^2 \rangle$$

We have direct measurements of $\sigma_C, \sigma_A,$ and r but $\langle \varepsilon^2 \rangle, \langle \eta^2 \rangle, \langle \zeta^2 \rangle, f$ are not known. However, we can deduce a few things just from the positivity of the variances:

1. $\langle \zeta^2 \rangle = \dfrac{\sigma_A^2 - \sigma_A \sigma_C r}{1-f} > 0$ implies $r < \dfrac{\sigma_A}{\sigma_C}$, thus a strict upper bound on *r*.
2. $\langle \eta^2 \rangle = \sigma_C^2 - \langle \varepsilon^2 \rangle = \sigma_C^2 - \dfrac{\sigma_A \sigma_C r}{f} > 0$ implies $f > \dfrac{\sigma_A}{\sigma_C} r$, thus a strict lower bound on *f*.

   The upper bound is, of course, $f = 1$.
3. Finally, for any given value of *f* we can solve for $\langle \varepsilon^2 \rangle, \langle \eta^2 \rangle,$ and $\langle \zeta^2 \rangle$ as follows:

$$<\varepsilon^2> = \frac{\sigma_A \sigma_C r}{f}$$
$$<\eta^2> = \sigma_C^2 - <\varepsilon^2>$$
$$<\zeta^2> = \frac{\sigma_A^2 - \sigma_A \sigma_C r}{1-f}$$

A graph of $\langle \varepsilon^2 \rangle, \langle \eta^2 \rangle,$ and $\langle \zeta^2 \rangle$ versus *f* for the setting phenomenon (type 3) is shown in Figure 1 (and the graphs for the other phenomena are similar). We see clearly that if one assumes a value of *f* near its lower bound then it follows that Ptolemy's measurements must have been *much* more accurate than Hipparchus', i.e. when *f* is near its lower bound, the implied value of $\langle \zeta^2 \rangle$, the variance in Ptolemy's measurements, quickly approaches zero. Probably the most likely scenario is that Hipparchus' lost catalog and the stars measured by Ptolemy have about the same variance, thus the fraction of stars copied in that scenario is given by the points where $\langle \varepsilon^2 \rangle = \langle \zeta^2 \rangle$, which implies $f = \frac{\sigma_C}{\sigma_A} r$, or f in the range 0.82-0.99, with the high end favored[9], as shown in the last column of Table 1. The net result is that, given the model assumptions, we have a quantitative estimate of the fraction of stars copied by Ptolemy, and that fraction is large.

This result should, of course, be kept in perspective. The most important limitation is that the Commentary refers only to the first and last stars to rise and set for each of 42 constellations, and not counting duplicates this amounts to just 134 stars. Thus strictly speaking, the model makes no assumption and offers no result on the many stars that appear in the Almagest but not in the Commentary. Still, it would be curious indeed if Ptolemy had happened to limit his copying to just those stars that are the first and last to rise and set in the various constellations. Further, the clear association between the *systematic* errors of the Commentary and the Almagest catalog established in Ref. 4 also implies that the copying was more broadly based, since the systematic errors are relatively smooth, few-parameter, collective effects that permeate the entire data sets in both the Commentary and the Almagest catalog.

	σ_A	σ_C	r	σ_A/σ_C	$f_{\min}$	$\sigma_C r / \sigma_A$
Λ_1	1.19	1.48	0.80	0.80	0.64	0.99
Λ_2	1.34	1.55	0.81	0.87	0.70	0.94
Λ_3	1.54	2.06	0.71	0.75	0.53	0.95
Λ_4	1.49	1.76	0.69	0.85	0.58	0.82

Table 1. Numerical results for the variances (columns 2-3), the correlation r (column 4), and the various derived quantities discussed in the text. The simple zero-parameter model for r predicts the results shown in column 5. In the extended model, column 5 is a strict upper bound on r. Column six shows the strict lower bound on the fraction of copied stars, while column 7 shows the fraction copied under the additional assumption that Hipparchus and Ptolemy measured star positions with the same accuracy.

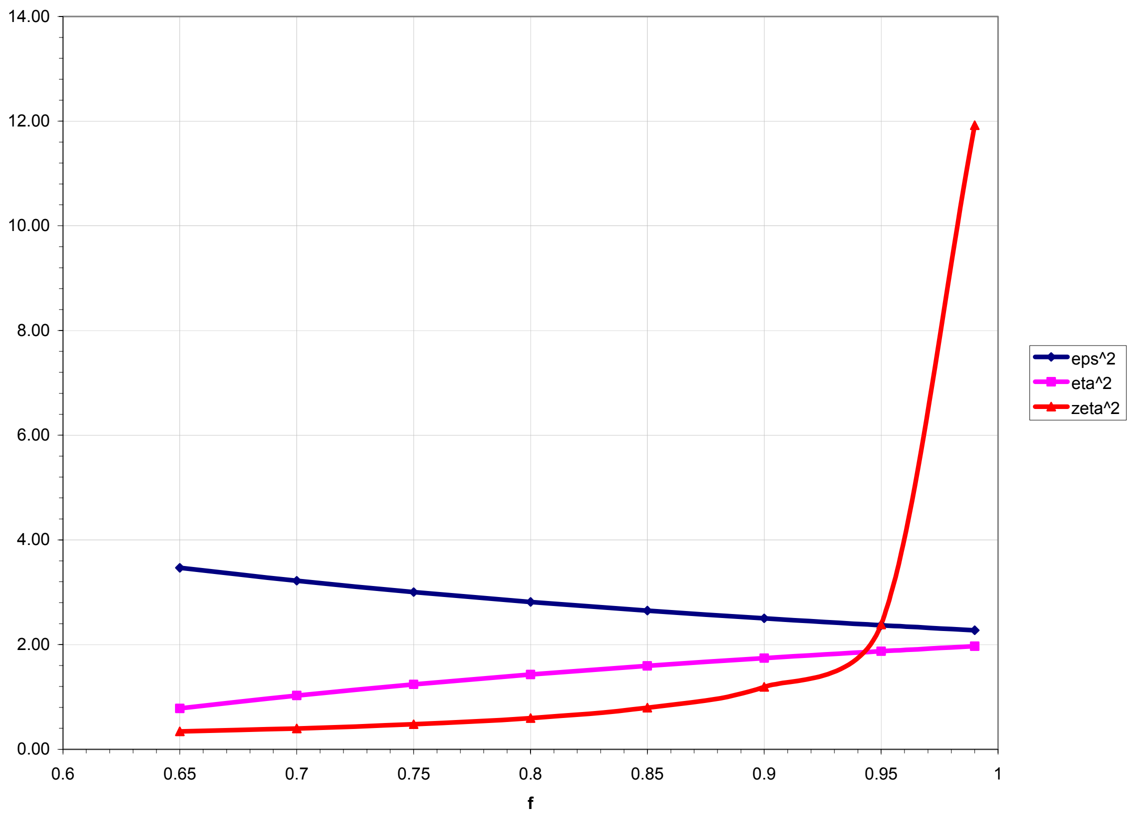

Figure 1. $\langle \varepsilon^2 \rangle, \langle \eta^2 \rangle$, and $\langle \zeta^2 \rangle$ (in degrees) versus f for the type 3 phenomenon. Note that as f drops below 0.95 then the variance of Ptolemy's measurements must get very small compared to the variance of Hipparchus' measurements.

[1] Hipparchus, *In Arati et Eudoxi phaenomena commentariorium*, ed. and transl. by K. Manitius (Leipzig, 1894).

[2] *Ptolemy's Almagest*, transl. by G. J. Toomer (London, 1984).

[3] H. Vogt, "Versuch einer Wiederstellung von Hipparchs Fixsternverzeichnis", *Astronomische Nachtrichten*, ccxxiv (1925), cols 2-54; G. Grasshoff, *The history of Ptolemy's star catalogue* (New York, 1990).

[4] D. Duke, "Associations between the ancient star catalogues", *Archive for history of the exact sciences* (forthcoming).

[5] Extensive simulation, discussed in Ref. 4, establishes the statistical strength of this claim.

[6] N. M. Swerdlow, "The enigma of Ptolemy's catalogue of stars", *Journal for the history of astronomy*, xxiii (1992), 173-183.

[7] J. Evans, "The Ptolemaic star catalogue", *Journal for the history of astronomy*, xxiii(1992), 64-68.

[8] James Evans (private communication).

[9] Although Hipparchus gives us, for each star, the degree of the ecliptic on the horizon and the degree of the ecliptic that is culminating, he is redundant in giving us both, since one is easily calculated from the other. Thus the results for the four types of phenomena are not independent, and the deviation shown in Table 1 for the fourth phenomenon is of little significance.