Bayesian inference for orbital eccentricities

L. B. Lucy

Astrophysics Group, Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2AZ, UK
e-mail: 1.lucy@imperial.ac.uk

Received 14 June 2012 / Accepted 21 December 2012

ABSTRACT

Highest posterior density intervals (HPDIs) are derived for the true eccentricities $e$ of spectroscopic binaries with measured values $e \approx 0$. These yield upper limits when $e$ is below the detection threshold $e_0$, and seamlessly transform to upper and lower bounds when $e > e_0$. In the main text, HPDIs are computed with an informative eccentricity prior representing orbital decay due to tidal dissipation. In an appendix, the corresponding HPDIs are computed with a uniform prior and are the basis for a revised version of the Lucy-Sweeney test, with the previous outcome $e = 0$ now replaced by an upper limit $e_U$. Sampling experiments with known prior confirm the validity of the HPDIs.

Key words. binaries: spectroscopic – methods: statistical – methods: data analysis

1. Introduction

For over two centuries, astronomers have been able to detect and analyse orbital motions for objects beyond the solar system. From measured positions (visual binaries) or radial velocities (spectroscopic binaries, exoplanets), orbital elements and their standard errors are typically obtained by least-squares. Accordingly, at this late date, when orbital motion is detected, prior information concerning similar objects is available and can be incorporated into the analysis.

An example is the Lucy-Sweeney (1971; LS) test for the statistical significance of small measured eccentricities $e$ for spectroscopic binaries (SB’s). The prior information that provided support for the LS test was as follows:

1) The small $e$’s (typically $\leq 0.05$) of numerous catalogued SB’s were $\sim E(e|0)$, the expected value of $e$ due to measurement errors when the true value is $e = 0$.

2) Savedoff’s (1951) investigation of $e \cos \omega$, where $\omega$ is the longitude of periastron. If an SB is also an eclipsing binary (EB), $e \cos \omega$ can be determined from the velocity curve (sp) and independently from the light curve (ph). For EB’s with secondary eclipses midway between consecutive primary eclipses, ($e \cos \omega)_p = 0$, but the values of ($e \cos \omega)_q$ are scattered over the interval $(-0.05, 0.05)$, confirming that non-zero $e$’s $\leq 0.05$ are often spurious.

3) Tidal dissipation gives $e \rightarrow 0$ as $t \rightarrow \infty$ since, for two point masses, $e = 0$ is the state of minimum orbital energy for fixed orbital angular momentum. Thus, if the time constant for this decay is short enough, the system is likely to be observed when $e \ll E(e|0)$ – i.e., well below the measurement threshold.

In view of this prior information, LS adopted $e = 0$ as the preferred (null) hypothesis ($H_0$) and imposed a moderately demanding level of significance before rejecting $H_0$ and accepting an elliptical orbit.

Given that secular evolution due to tidal dissipation was already well established in 1971 and is not less so now, there is merit in explicitly incorporating this mechanism into the analysis rather than implicitly via the LS preference for $e = 0$. This can be achieved by replacing the frequentist approach of LS by one based on Bayes’ theorem.

2. Estimating the true eccentricity $e$

We suppose that radial velocities of a single-lined spectroscopic binary (SB1) have been analysed to estimate the orbital elements and their standard errors. We ask: what can be inferred about the error-free elements when orbital evolution due to tidal dissipation is taken into account?

2.1. Posterior probability

We adopt the notation used in Lucy (1974; L74). The vectors of the estimated and the error-free elements are denoted by $x$ and $\xi$, respectively; and the distribution of probability in x-space for given $\xi$ is denoted by $P(x|\xi)dx$. Integrating over $\xi$-space, we find that the distribution of probability in $\xi$-space is $\phi(x)dx$, where

$$\phi(x) = \int P(x|\xi)\Pi(\xi) d\xi.$$  \hspace{1cm} (1)

Here $\psi$ is the probability density function (pdf) that represents our prior knowledge about $\phi(\xi)d\xi$, the distribution of probability in $\xi$-space for the SB1’s true elements.

Since SB1’s with small $e$’s are of interest, we suppose that Sterne’s (1941) elements have been chosen, as in LS. The six elements are then: $P$, the orbital period; $\gamma$, the systemic velocity; $K$, the semi-amplitude of the velocity curve; $T_0$, an epoch at which the mean longitude is zero; and the pair $e \cos \omega$, $e \sin \omega$. In terms of these parameters, the radial velocity curve of an SB1 is given by Eq. (1) in LS.

With this choice and the assumption $e^2 \ll 1$, the off-diagonal elements of the least-squares matrix have zero expectation values when observational weight is uniformly distributed in phase (LS). Since observers strive to meet this condition, we assume...
it to be true; and this then implies negligible correlations between the elements. Accordingly, to a good approximation, the error-broadening kernel \( \Pi(\xi, \eta) \) is simply the product of the six Gaussians giving the independent error distributions of the six elements.

Next consider the pdf \( \psi(\xi, \eta) \), which is convolved in Eq. (1) with the kernel \( \Pi(\xi, \eta) \). It follows that \( \psi \) is only relevant if it varies significantly within the error bars of an individual orbital element. This is not true for \( P, \gamma, K, \) and \( T_0 \). However, when orbital circularisation is taken into account, \( \psi(\xi, \eta) \) may vary significantly within the domain \( (e \cos \sigma, e \sin \sigma, \epsilon \sin \omega, \epsilon \cos \omega, e \sin \omega, \epsilon \cos \omega) \).

Accordingly, we now integrate Eq. (1) with respect to \((P, \gamma, K, T_0)\)-components of \( x \) and use the normalization of \( \Pi(\xi, \eta) \) to respect to each of these four elements. The result is

\[
\phi_\epsilon(x, y) = \int \psi_\epsilon(\xi, \eta)\Pi_\epsilon(x - \xi, y - \eta) d\xi d\eta
\]

where \( x = e \cos \omega \) and \( y = e \sin \omega \) are the two surviving estimated quantities, whose error-free values are \( \xi = e \cos \sigma \) and \( \eta = e \sin \sigma \), respectively, and the subscript \( \epsilon \) indicates that the pdf’s are defined on a Cartesian grid.

We now assume that the standard errors of \( e \cos \omega \) and \( e \sin \omega \) are equal, which is true of their expectation values when observational weight is uniformly distributed in phase (LS). The broadening kernel is then the circular normal distribution

\[
\Pi_\epsilon(x - \xi, y - \eta) = \frac{1}{2\pi\mu^2} \exp \left[ -\frac{(x - \xi)^2 + (y - \eta)^2}{2\mu^2} \right]
\]

where \( \mu = \sigma_{e \cos \omega} = \sigma_{e \sin \omega} \).

Bayes’ theorem can now be invoked to derive – see Eq. (10) in L74 – the posterior probability

\[
Q_\epsilon(x, y) = \frac{\psi_\epsilon(\xi, \eta)\Pi_\epsilon(x - \xi, y - \eta)}{\psi_\epsilon(x, y)}
\]

Thus, if the least-square solution is

\[
x = e \cos \omega \pm \mu, \quad y = e \sin \omega \pm \mu
\]

the distribution of probability in \( (e \cos \omega, e \sin \omega) \)-space is given by \( Q_\epsilon(x, y) = Q_\epsilon(x, y) \Pi_\epsilon(x - \xi, y - \eta) \).

For the problem at hand, polar \( (p, \varphi) \) coordinates \( (\xi, \eta) \) are more convenient than the Cartesian coordinates \( (x, y) \) in the above formulae. The Jacobian of the transformation is \( J = e \), so that, by conservation of probability,

\[
\psi_p(\epsilon, \varphi) = e \psi_\epsilon(\xi, \eta)
\]

and

\[
Q_p(\epsilon, \varphi|x, y) = e Q_\epsilon(\xi, \eta|x, y).
\]

2.2. A physical model for \( \psi_p \)

Given \( P \), there is a 2D family of binaries that match the measured \( K \) at some inclination. For simplicity, a representative example is chosen to avoid integrating over all possibilities.

Now consider an ensemble of such binaries that form at a uniform rate in the solar neighbourhood and have lifetime \( t_\epsilon \). We further suppose that all have \( \epsilon = \epsilon_0 \) at \( t = 0 \) and that thereafter \( \epsilon \) decays exponentially with fixed \( \epsilon \)-folding time \( t_\epsilon /\nu \), so that

\[
\ln \epsilon(t) = \ln \epsilon_0 - \nu \frac{t}{t_\epsilon}
\]
3.1. Parameters

There are two basic parameters, $e/\mu$ and $v$, the number of $e$-folding decay times in $t_e$.

Note that $e_0$ is not a consequential parameter provided that $e \ll e_0$. In effect, we assume that an SB1 with $e = 0$ has reached this configuration due to secular evolution and not due to the formation mechanism. In these calculations, $e_0 = 0.5$.

We choose $v = 8.52$, so that $e_0 = 10^{-4}$. Then, with $\mu = 0.01$, $e(t) < 2.45\mu$, the LS threshold, when $t > 0.35t_e$. Thus, from our ensemble of SB1's, $e=0$ would be assigned $e = 0$ by the LS test.

3.2. The posterior pdf $\chi(\log e|e)$

Because of the concentration of probability towards $e_*$, plots are more informative if the abscissa is $\log e$ rather than $e$. Accordingly, we define

$$\chi(\log e|e) = e q(e|e) \times \ln 10.$$  \hspace{1cm} (13)

In Fig. 1, this pdf is plotted for $e/\mu = 1.0, 2.45, 3.03$ and $3.72$, values selected as follows: if $e = 0$, LS showed that the probability $p_0$ of exceeding $e$ is given by

$$\ln p_0 = \frac{1}{2} \left( \frac{e}{\mu} \right)^2$$  \hspace{1cm} (14)

provided that $\mu \ll 1$. Therefore, when testing $H_0$, the above values of $e/\mu$ correspond to levels of significance 61, 5, 1 and 0.1%, respectively. The criterion $e/\mu > 1$ was proposed and implemented by Luyten (1936); the 5% level by LS.

For $e/\mu = 3.72$, $\chi$ is an asymmetric bell-shaped function peaking at $e_0$, but with a tail extending down to $\log e = -4.0$. As $e/\mu$ decreases, the peak weakens and the tail strengthens. At $e/\mu = 1.0$, the peak is absent and all the probability is in the tail, which derives from the physical model. Intermediate calculations show that the peak first appears at $e/\mu = 1.42$. Thus, for $e/\mu < 1.42$, $\chi$ is a monotonically decreasing function of $e$. For $e/\mu > 1.42$, $\chi$ is unimodal.

The pdf $\chi$ for $e/\mu = 3.72$ and $\mu = 0.01$ in Fig. 1 is computed with $e_0/\mu = 50$. Repeating this calculation shows that $\chi$ is independent of the upper limit $e_0$ provided that $e_0/\mu \geq 10$.

3.3. Percentiles

The posterior pdf’s in Fig. 1 imply asymmetric and rapidly changing credible (or Bayesian confidence) intervals as $e/\mu$ varies. These are plotted in Fig. 2 for the indicated values of the probability that the true eccentricity is $<e$. For the normal distribution, the values 0.159, 0.500, 0.841 and 0.977 correspond to displacements of $-1, 0, +1$ and $+2\sigma$, respectively.

Figure 2 reveals a dramatic switch in solution regime at $e_1/\mu \approx 3.6$. For $e \geq e_1$, the measured value $e$ is close to the 50% percentile and is tightly enclosed by the $\pm 1\sigma$ intervals. In this regime, inferences are dominated by the actual measurement $e \pm \mu$. But this ceases to be so for $e \leq e_1$. Thus, for $e/\mu < 2.65$, the “solution” $e = \bar{e}$ falls outside the $\pm 1\sigma$ intervals. Evidently, for $e \leq e_1$, inferences are increasingly dominated by the model of Sect. 2.2.

3.4. Highest posterior density intervals

For $e/\mu < 1.42$, $\chi$ is a monotonically decreasing function of $e$ (Sect. 3.2). It follows that traditional, equal-tail credible intervals exclude the point $(e = e_*)$ with greatest probability density (pd). This undesirable feature is avoided by instead computing highest posterior density intervals (HPDI; Box & Tiao 1973). These intervals are such that every point included has a higher pd than every point excluded.
In general, the calculation of HPDIs is non-trivial. But here the pdf’s χ are not pathological (Fig. 1), and so the following clipping algorithm finds the HPDI for specified ε and designated enclosed probability (1 − α).

Let log ε_k, ..., log ε_k be consecutive grid points that belong to and define the HPDI (ε_k, ε_k). Then an HPDI with smaller included probability is obtained by eliminating the grid point log ε_k if χ_k < χ_k or the grid point log ε_k if χ_k > χ_k. This is repeated until the included probability = (1 − α).

The 95% HPDIs thus obtained for χ(log ε) are plotted for ε = 0.000(0.001)0.080 in Fig. 3. For ε < 0.018, the HPDIs are effectively one-tail intervals since the lower bound is ε_L = ε, = 10^{-4}. Accordingly, for this problem, HPDIs provide a seamless transition from upper limits for non-detected to two-sided intervals for detected eccentricities (cf. Feldman & Cousins 1998). This is an appealing aspect of HPDIs for interpreting measured ε’s and their uncertainties.

For ε < 0.0142, the interval excluded from an HPDI is a single-tail because of the aforementioned monotonicity. For ε > 0.0142, the pdf χ is unimodal (Fig. 1), with a maximum whose location → ε as ε/μ increases. This emerging, measurement-driven maximum eventually brings about the transition from one- to two-tailed intervals. For ε > 0.018, the excluded probability is contained in two tails, with the upper tail’s probability being initially 5%, but this decreases to ≈2.5% when ε/μ ≈ 1 because of χ’s increasing symmetry – see Fig. 1.

Note that for the normal distribution, Ν(0, 1), the 95% HPDI is the familiar equal tail interval (−1.96, +1.96) and the width of this interval is the narrowest that encloses 95% of the probability. For non-symmetric pdf’s, HPDIs are the narrowest intervals enclosing probability (1 − α) and as such are a natural generalization of the conventional equal-tail intervals for symmetric, bell-shaped pdf’s.

Confidence intervals are an economical means of conveying the compactness or otherwise of a variate’s distribution. The resulting loss of information, if of concern, can be avoided by plotting the pdf’s, as in Figs. 1 and A.2.

### 3.5. Detection threshold

The transition from one- to two-tailed HPDIs is a natural definition of the detection threshold ε_L for non-zero eccentricity – i.e., the measured value ε above which attribution purely to measurement errors is implausible. However, Fig. 3 shows that ε_L remains ≪ε for a considerable interval beyond ε_L = 0.018, which in any case depends on ν, a parameter likely to be only crudely estimated.

If detection is crucial for a subsequent investigation – e.g., an observing program – then a threshold closer to ε should be adopted (Figs. 2 and 3).

### 3.6. Simulation

The role that HPDIs can play in reporting eccentricities is best illustrated by sampling experiments. Accordingly, synthetic data for the model of Sect. 2.2 are created as follows: if z_1, z_2 are random numbers in (0, 1), a random ensemble member in (ε, ν)-space is at

\[ ε = e_0 \exp(-νz_1) \quad \nu = 2\pi z_2. \]  \hfill (15)

Then, if ε_1, ε_2 are random Gaussian variates, this ensemble member is observed at the point

\[ e \cos \omega = e \cos \nu + με_1 \quad e \sin \omega = e \sin \nu + με_2. \]  \hfill (16)

Repeated N times, the resulting ε’s comprise a simulated observing campaign of N random ensemble members whose exact eccentricities ε are known.

With ε_L = 10^{-4}, a sample of N = 1000 SB1’s are plotted in Fig. 4. As expected, the large majority of the points fall within the 95% HPDIs (ε_L, ε_L). With Luyten’s criterion ε > μ, 79.5% of this sample would have their elliptical orbits accepted. But Fig. 4 clearly shows that most systems with ε ∈ (0.01, 0.03) have
\( \varepsilon < 0.01 \) and so exceed Luyten’s criterion because of the bias of the non-negative estimator \( e \) (LS). With the LS criterion \( e > 2.45 \mu \), the accepted percentage drops to 40.5, and most of the systems with \( e \in (0.01, 0.03) \) would now be assigned \( e = 0 \), a marked improvement.

The further improvement provided by the HPDIs is that the assignment \( e = 0 \) can now be replaced by an upper limit. Thus, for example, with this choice of prior, an SB1 with marked improvement.

### Table 1. HPDI 95% upper limits \( e_{95} \) when \( e/\mu = 1.0 \).

| \( \log e \) | \( \nu \) | \( e_{95}/\mu \) |
|------------|------|--------------|
| -3.0       | 6.21 | 1.71         |
| -4.0       | 8.52 | 1.37         |
| -5.0       | 10.82| 1.15         |
| -6.0       | 13.12| 0.98         |
| -7.0       | 15.42| 0.85         |
| -8.0       | 17.73| 0.74         |
| -9.0       | 20.03| 0.65         |

For \( \nu \), the non-negative estimator \( e \) (LS). With the LS criterion \( e > 2.45 \mu \), the accepted percentage drops to 40.5, and most of the systems with \( e \in (0.01, 0.03) \) would now be assigned \( e = 0 \), a marked improvement.

Thus, for example, with this choice of prior, an SB1 with marked improvement.

### Appendix A: Uniform prior

The model of Sect. 2.2 is not appropriate if the SB or star-planet system has additional components causing significant gravitational perturbations. In this circumstance, a sensible option is to assume a uniform prior for \( e \), as is already common practice for exoplanets (e.g., Ford 2006; Eastman et al. 2013).

### Conclusion

In this paper, by incorporating a model of an SB1’s secular evolution, Bayes’ theorem is used to infer bounds on its exact eccentricity \( e \) given its measured value \( e + \mu \). Because the system’s lifetime \( \tau \) is finite, the asymptote \( e = 0 \) is never reached. Thus, in contrast to Luyten (1936) and LS, the statistical problem is not one of model selection. Systems assigned \( e = 0 \) by these earlier tests should preferably have upper limits \( e_{U} \) computed.

As Fig. 4 shows, the Bayesian upper limits contain the systems for which \( e \) is significantly larger than \( e \) due to measurement errors and bias. Thus, a major historical cause of spurious \( e \)’s is eliminated. But physical causes remain, such as those due to proximity effects or to additional line absorption by gas streams. An example is TR A, for which Skuljan et al. (2004) improved the precision of the radial velocities by a remarkable factor of 100 and reported a small but highly significant

\[ e = 0.0140 \pm 0.0002. \] However, the significant non-detection of the Keplerian third harmonic (Lucy 2005) invalidated this claim. As precision improves, similar testing for the third and higher Keplerian harmonics is essential for confirming that an orbit is truly eccentric. In addition, an update of Savedoff’s (1951) work would provide numerous examples of spurious \( e \)’s for investigation into physical causes other than measurement bias.

### Acknowledgements

I am grateful to the referee for pointing out an error in statistical terminology.

### Table A.1. Detection thresholds for \( e \).

| \( \alpha(\%) \) | \( e_{0.5}/\mu \) | \( e_{0.08}/\mu \) |
|-------------|----------------|----------------|
| 31.7        | 1.52           | 1.71           |
| 10.0        | 2.15           | 2.21           |
| 5.0         | 2.45           | 2.49           |
| 1.0         | 3.03           | 3.06           |
| 0.1         | 3.72           | 3.74           |

1. Hearnshaw et al. (2012) have just reported eleven additional nondetections.
The enclosed posterior probability of each HPDI is 95%, and the prior pdf $q(\omega | \psi) \propto \exp(-\omega^2/2\mu^2)$ is given by Eq. (A.1). The filled circles are the posterior means $\langle \omega \rangle$ computed from Eq. (12).

Table A.2. Upper limits (bounds) for $\epsilon$.

| $e/\mu$ | $e_{0.024}/\mu$ | $e_{0.037}/\mu$ | $e_{0.073}/\mu$ | $e_{0.094}/\mu$ | $e_{0.094}/\mu$ |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0.0     | 1.00            | 1.64            | 1.96            | 2.58            | 3.29            |
| 0.5     | 1.06            | 1.75            | 2.08            | 2.73            | 3.48            |
| 1.0     | 1.28            | 2.05            | 2.41            | 3.11            | 3.89            |
| 1.5     | 1.66            | 2.50            | 2.88            | 3.59            | 4.38            |
| 2.0     | (2.59)          | 3.03            | (3.41)          | 4.12            | 4.90            |
| 2.5     | (3.31)          | (3.87)          | (3.96)          | 4.65            | 5.43            |
| 3.0     | (3.86)          | (4.52)          | (4.82)          | 5.18            | 5.96            |
| 3.5     | (4.37)          | (5.04)          | (5.36)          | (5.97)          | 6.48            |
| 4.0     | (4.89)          | (5.54)          | (5.86)          | (6.49)          | (7.18)          |

The revised LS test with $1 - \alpha = 0.95$ proceeds as follows:

1) The eccentricity $\epsilon \pm \mu$ is derived from the least squares solution.
2) If $e/\mu > 2.49$, this measured value $\epsilon \pm \mu$ is accepted.
3) However, if $e/\mu < 2.49$, the measured value is replaced by the upper limit $e_{0.037}$ obtained by interpolation in Col. 4 of Table A.2.

To illustrate this revised test, upper limits are given in Table A.3 for six SB1’s for which circular orbits are reported on the first page of Table I in LS. Thus for YZ Cas, the least squares value $\epsilon = 0.004$ was rejected by the LS test and $\epsilon = 0$ accepted. We now compute the 95% upper limit as follows: the estimate $\mu = 0.0037$ is derived from Eq. (13) and Table I of LS. Linear interpolation in Table A.2 at $e/\mu = 1.1$ then gives $e_{0.037}/\mu = 2.50$, whence $\epsilon < e_{0.037} = 0.009$.

Table A.3. Upper limits $e_{0.037}$ for SB1 sample.

| Star     | $e$ | $\mu$ | $e/\mu$ | $e_{0.037}/\mu$ | $e_{0.037}/\mu$ |
|----------|-----|-------|--------|-----------------|-----------------|
| YZ Cas   | (0.004) | 0.0037 | 1.10   | 2.50             | <0.009          |
| HD 7345  | (0.046) | 0.019  | 2.47   | 3.93             | <0.073          |
| DM Per   | (0.09)  | 0.074  | 1.22   | 2.62             | <0.19           |
| HD 16589 | (0.008) | 0.0087 | 0.92   | 2.35             | <0.021          |
| HD 18337 | (0.073) | 0.030  | 2.48   | 3.93             | <0.12           |
| HD 21912 | (0.005) | 0.011  | 0.45   | 2.07             | <0.023          |

The neutral standpoint is an attractive basis for a revised version of the LS test in which the previous acceptance of a circular orbit $(H_0)$ is now replaced by an upper limit.

Fig. A.1. Highest posterior density intervals for the posterior pdf $q(\omega | \psi)$. The hatched area with $\epsilon > 0.0249$, the measured value for marginal detection (Sect. 3.5). The hatched area with $\epsilon > 0.0394$ contains 5% of the probability. Note that $q(\omega | \psi) = q(\omega | H_0)$.

Fig. A.2. Detection threshold. The posterior pdf $q(\epsilon | \psi)$ for $e_{0.0249} = 0.0249$, the measured value for marginal detection (Sect. 3.5). The hatched area with $\epsilon > 0.0394$ contains 5% of the probability. Note that $q(\omega | \psi) = q(\omega | H_0)$.

Table A.3. Upper limits $e_{0.037}$ for SB1 sample.