# When time matters

D. Pourbaix
*FNRS, IAA, Université Libre de Bruxelles, 1050 Brussels, Belgium. pourbaix@astro.ulb.ac.be*

## Introduction

Whereas deriving the orbit of a binary star was perceived as a tedious task even during the second half of the 20th century, nowadays, computers can investigate millions of solutions and keep the best one in just a few seconds. However, even if one is very careful in defining what best means (e.g. least-squares sense), the nonlinearity of the model can be very misleading, thus possibly yielding some solutions which are numerically valid but physically rather unlikely. With this paper, we aim at warning the fitters about the potential inadequacy of the methods they use to assess their solutions. Indeed, the problem is neither in the model nor in the fitting method but rather in the often overlooked assumptions of some statistical tools.

## 1. Model assessment

### 1.1 *Linear fit*

Under the assumption of Gaussian observational errors, the least-squares paradigm ($L2$-norm), combined with a good minimisation scheme (QR, Levenberg-Marquardt), is often adopted to fit these observations with some model. The value of the objective function (usually abusively called $\chi^2$) together with the number of data points ($N$) and the number of parameters ($p$) in the model allow the end-user to assess how good the fit is thanks to some auxilliary tables of statistical quantities.
Alternatively, the value of the objective function can be combined with the numbers of observations and of parameters to yield a number from which it is easier to apprehend the quality of the fit directly. An example of such combination is the goodness of fit, $F2$, used to assess the quality of the fits in the Hipparcos catalogue [5]. It is defined as

$$F2 = \sqrt{\frac{9\nu}{2}}\left(\sqrt[3]{\frac{\chi^2}{\nu}} + \frac{2}{9\nu} - 1\right) \tag{1}$$

where $\nu$ is the number of degrees of freedom (i.e. $N-p$) and $\chi^2$ is the weighted sum of the squares of the differences between the predicted and the observed positions [10]. $F2$ (Wilson & Hilferty's cube root transformation) follows approximately a $N(0,1)$ distribution [15]. The behaviour of the Hipparcos single stars (i.e. a linear model with 5 parameters) is very close to the theoretical prediction. Their goodness of fit follows a $N(0.22, 1.08)$ distribution.
In front of two alternative models with respectively $p$ and $p'$ parameters ($p' > p$), one often relies upon an $F$-test [2] to decide whether the reduction of the objective function achieved with the additional parameter(s) is significant or not.

$$\hat{F} = \frac{N-p'}{p'-p}\frac{\chi^2_p - \chi^2_{p'}}{\chi^2_{p'}} \tag{2}$$

$$\alpha = Pr\left[\hat{F} < F(p'-p, N-p')|\text{no signal}\right] \tag{3}$$

where $\alpha$ is related to the rate of false detection. If the first linear model is already good and the second model is linear as well, comparing their respective $F2$ does not help as they would essentially be the same, thus concluding that both models are good. Instead, in such a case, the $F$-test will notice that the reduction of the $\chi^2$ is consistent with the number of additional parameters and the new model discarded.

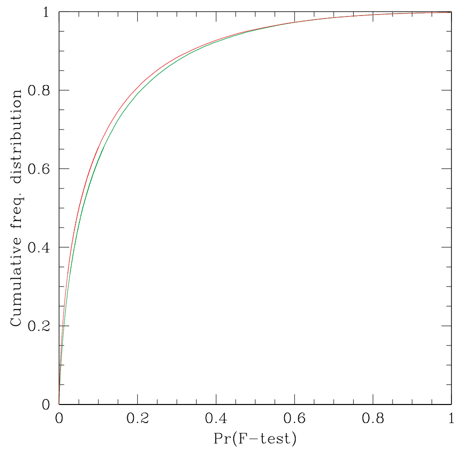


**Figure 1:** Fraction of single stars for which the orbital model is accepted. When the acceleration model is tested on the same sample, one obtains the diagonal ((0,0)–(1,1)). The lower line corresponds to a sample of synthetic single stars whereas the red one is based on the Hipparcos single star observations.

Once again, when an acceleration model (i.e. seven parameters) is applied to the Hipparcos single stars (i.e. those for which the five parameter model was already acceptable), the percentage of false detection perfectly follows the threshold set on $\alpha$. So, nothing to worry about, if one tries a model too sophisticate, the $F$-test concludes that it does not improve over the simpler model.

### 1.2 *Nonlinear models*

Can the whole previous section be applied to any nonlinear model? What if one applies the orbital model (a seven parameter model, nonlinear, on top of the five parameter, linear, single star one)? The behaviour of the F-test comparing the full orbital model with the single star one is depicted in Fig. 1. For a substantial fraction of the thought-to-be single stars, the reduction of the objective function achieved with the orbital model is considered significant. Even if one cannot reject that it is true for some of them, their number is rather suspicious: at a 1% false detection level, more than 21,400 objects out of 100,000 single stars would turn out to be binaries. A similar experiment on synthetic single stars yields 17,400 binaries.

Why are there so many accepted binaries? Let $\mathbf{v}$ denotes the observations, $\mathbf{a}$ the $p$ parameters, $\mathbf{t}$ the $N$ observation times, $\mathbf{V}$ the weighting matrix and $f$ the model:

$$\Xi^2 = (\mathbf{v} - f(\mathbf{a},\mathbf{t}))^t \mathbf{V}^{-1} (\mathbf{v} - f(\mathbf{a},\mathbf{t})) \sim \chi^2_{N-p} \tag{4}$$

is true if and only if $f$ is linear. In the case of the orbital model, $N - 12$ therefore overestimates the number of degrees of freedom.

## 2. Non-parametric tests

Unlike the F-test, the Kolmogorov-Smirnov (KS) test does not rely upon any assumption about the very nature of the model (non parametric test). It is used, for instance, to assess the confidence level at which one can reject the hypothesis that two empirical distributions are drawn from the same parent population. If one denotes by $S_1$ and $S_2$ the cumulative distributions of the two populations, the Kolmogorov-Smirnov statistic is:

$$D_{\mathrm{KS}} = \max_{-\infty<x<\infty} |S_1(x) - S_2(x)| \tag{5}$$

In the present case, the two distributions can be the residuals based on the single star solution and those derived with the orbital model.
The sets of signed weighted residuals (based on a least squares fit) are considered indistinguishable too often (thick line in Fig. 2). This is a pitfall of the KS test. Its sensitivity is not uniform over $x$. KS test is actually more suited to discriminate upon the medians of the distributions [9, 14, 1]. The median of the set of residuals is essentially zero in both cases and that is why KS does not distinguish the two. As an alternative, one can take the absolute value of those residuals but it affects the median so much that the test becomes over sensitive (line of medium thickness in Fig. 2).
Kuiper [11] proposed a statistic which overcomes the Kolmogorov-Smirnov maximum sensitivity at the median. The expression in Eq. 5 is replaced with

$$D_{\mathrm{Ku}} = \max_{-\infty<x<\infty}(S_1(x) - S_2(x)) + \max_{-\infty<x<\infty}(S_2(x) - S_1(x)) \tag{6}$$

The corresponding probabilities are plotted in Fig. 2 as a thin line just below the diagonal. Though this new test slightly underestimates the number of systems worth keeping (721 at 1% on synthetic data), its overall behavior is also the closest to the expectations (the cumulative frequency distribution of the probabilities is essentially parallel to the diagonal).

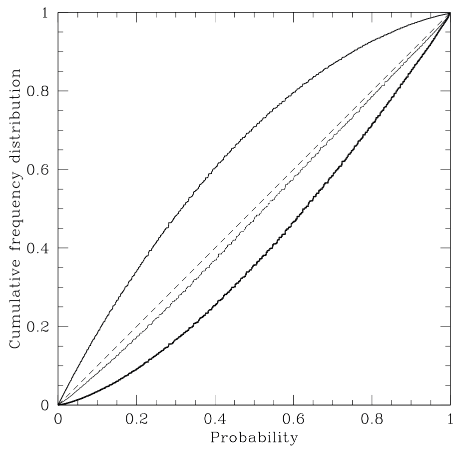


**Figure 2:** Cumulative frequency distribution of the Kolmogorov-Smirnov probabilities that the signed (resp. unsigned) residuals of the orbital and single star model are distinct drawn as a thick (resp. medium) line. The distribution of the Kuiper probabilities is plotted as a thin line just below the diagonal.

## 3. Back to F-test

The only reason for the hypothesis about the linearity of the model in the definition of $\chi^2$ is to make it straightforward to obtain the number of degrees of freedom. Very few references actually deal with the nonlinear case and, among them, the usual approach consists in correcting, say, the F-statistic to account for the nonlinearity by deriving the curvature of the model near the solution [3, 7, 4]. A truncated Taylor expansion of the model is also an alternative [6].
What we are still missing is a way to estimate the number of degrees of freedom in case of a nonlinear model or the smallest linear model (in terms of number of parameters) that yields the same level of $\Xi^2$ as the nonlinear model. Among linear expansions, Fourier sounds the most appropriate owing to the periodic nature of the nonlinear model.
Monet [13] used the Fourier transform to derive the orbital parameters of binaries by exactly solving for them from the minimum number of expansion coefficients. Such a minimalist approach behaves well

only for small eccentricities. However, we are not looking for the orbital solution, so there is no need to limit the expansion to low harmonic ranks.

Jarnadin [8] showed that the coefficient of each additional harmonic is dominated by a higher power of the eccentricity. Since successively higher powers drop off more slowly as the eccentricity approaches unity, the number of harmonics required for an adequate linear approximation increases with the eccentricity. Moreover, the value of $\Xi^2$ (Eq. 4) based upon that approximation also decreases as the number of harmonics grows. Let $\Xi_L^2$ be $\Xi^2$ based upon such a linear approximation with $L$ coefficients. We define the number of degrees of freedom of the nonlinear model as the smallest $L$ such that $\Xi_L^2 < \Xi^2$, supplemented by three to account for $e$, $P$, and the periastron time.

One could object that the coefficients in the Fourier expansion are not independent. Even though there are relations between them, these coefficients are linearly independent. That is the reason why additional terms (besides those listed by Lucy [12]) are required when dealing with larger eccentricities. Regardless of any functional dependency, the fact that all the coefficients are linearly independent is all it matters as far as the number of degrees of freedom is concerned. As the power of the leading term increases, for any eccentricity, there exists a coefficient index above which the coefficients are numerically linearly dependent.

One needs to be a bit careful about what is meant by linear approximation of the model and Fourier transform, especially in the context of astrometric binaries. The whole picture is a dual-period model: a one-year period describing the absolute position of the center of mass (initial position, parallax and proper motion) and the actual orbital model. The former is already linear so only the latter needs to be linearly approximated through the Fourier expansion. The two components (in $x$ and $y$-direction) corresponding to the time independent part of that expansion are embedded in the position of the center of mass so the expansion actually begins at the fundamental frequency.

As foreseen, the dimension of the linear approximation of the model increases with $e$. However, depending on the distribution of the observations, the very same eccentric orbit can be approximated with few or many Fourier terms. There is therefore no way to directly estimate $L$ from $e$ without evaluating several linear approximations. Owing to the speed of the step, this is a very minor drawback.

The reduction of the number of accepted ($\alpha = 1\%$) orbital solutions for synthetic single stars is substantial, from 17.3% to 0.7%. This is slightly below the theoretical value (only 1% of false detection is expected). Once applied to the original observations, the same procedure yields 1.3% of potential binaries. Only 48 objects are flagged as binaries with both data sets, thus suggesting that the distribution of the observing time might be responsible for the result in very few cases only.

## Conclusion

We have presented several approaches to circumvent the limitations of the F-test in case of a nonlinear model. However, it was somehow assumed that the orbital model was fitted right away, i.e. adding seven parameters to the single star model in once. In the framework of Gaia, the approach is slightly different: a cascade of linear models with increasing size is considered. Simulations show that with a false detection rate set to one per one thousand, we get four single stars with an orbital model out of 10,000.

## References


[1] Babu G. J. and Feigelson E. D. 1996. *Astrostatistics*. Chapman & Hall.

[2] Bevington P. R. and Robinson D. K. 1992. *Data reduction and error analysis for the physical sciences*. McGraw-Hill, 2nd edition.

[3] Bates D. M. and Watts D. G. 1980. Relative curvature measures of nonlinearity. *J. Royal Statistical Society*, B42:1–25.

[4] Bates D. M. and Watts D. G. 1988. *Nonlinear Regression Analysis & its Applications*. Wiley Interscience.

[5] ESA 1997. *The Hipparcos and Tycho Catalogues*. ESA SP-1200.

[6] Gallant A. R. 1977. Testing a nonlinear regression specification: a nonregular case. *J. Am. Statistical Association*, 72:523–530.

[7] Hamilton D. and Wiens D. 1987. Correction factors for f ratios in nonlinear regression. *Biometrika*, 74:423–425.

[8] Jarnadin M. P. Jr. 1965. Expansions in elliptic motion. *Astronomical Papers*, 18:1.

[9] Kendall M. G. and Stuart A. 1978. *The Advanced Theory of Statistics*. Griffin.

[10] Kovalevsky J. and Seidelmann P. K. 2004. *Fundamentals of Astrometry*. Cambridge University Press.

[11] Kuiper N. H. 1962. *Proceedings of the Koninklijke Nederlands Akademie van Wetenschappen Ser. A*, 63:38–47.

[12] Lucy L. B. 2005. Spectroscopic binaries with elliptical orbits. *A&A*, 439:663.

[13] Monet D. G. 1979. A method for solving binary star orbits using the Fourier transform. *ApJ*, 234:275.

[14] Press W. H., Teukolsky S. A., Vetterling W. T. , and Flannery B. P. 1992. *Numerical Recipes in C*. Cambridge University Press, 2nd edition.

[15] Stuart A. and Ord K. 1994. *Kendall's Advanced Theory of Statistics*, volume 1. Hodder Arnold, 6 edition.