Challenges in Facing Hundreds of Millions of Stars

D. Pourbaix

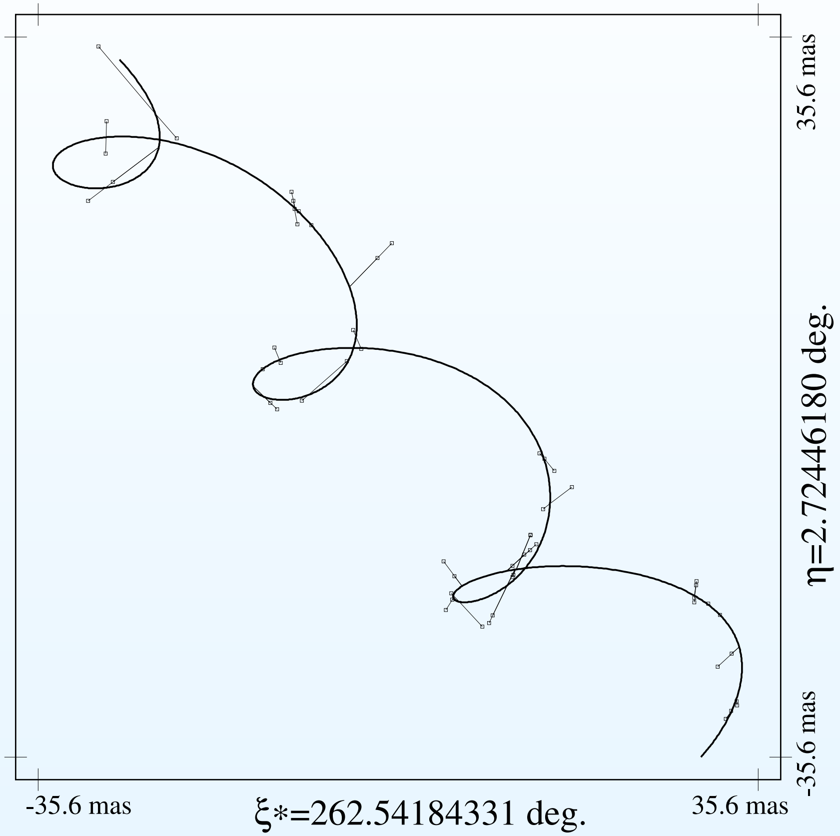
FNRS, IAA, ULB Brussels

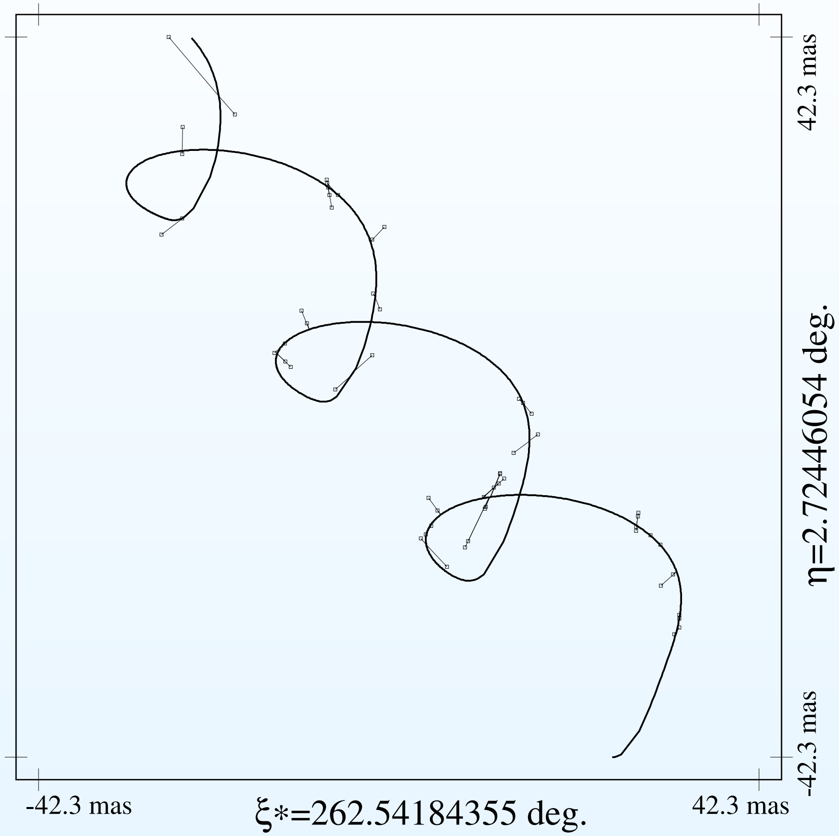
Collaborator: H.C. Smith Jr (University of Florida)

Overview

1. Visual, astrometric binaries and extrasolar astrometric orbit fitting
2. Spectroscopic watch dog
3. Blind fit
   - A priori versus a posteriori tests
   - Hipparcos as a testbed
   - Effect of nonlinearity
   - Alternative tests
   - Benefit from combining
   - Alternative explanations
   - Conclusion

Warning case





Two orbits of HIP 85749 with $e = 0.2$ and $e = 0.998$ (Pourbaix 2002).

1s per Gaia star: 38 years!

Core task – shell task

Core task (e.g. GIS) ⟷ Shell task (e.g. ABS module)

Criterion?

The optimal criterion is a compromise between:

- maximizing the scientific return (e.g. avoiding HD 209458-like situations with Hipparcos),
- minimizing the computing time ($60\,10^6$ AB would take 200 days on today machines with the present code).

A priori versus *a posteriori* assessment:

a priori test: there is no need to call the shell task (i.e. the single star fit is good enough);

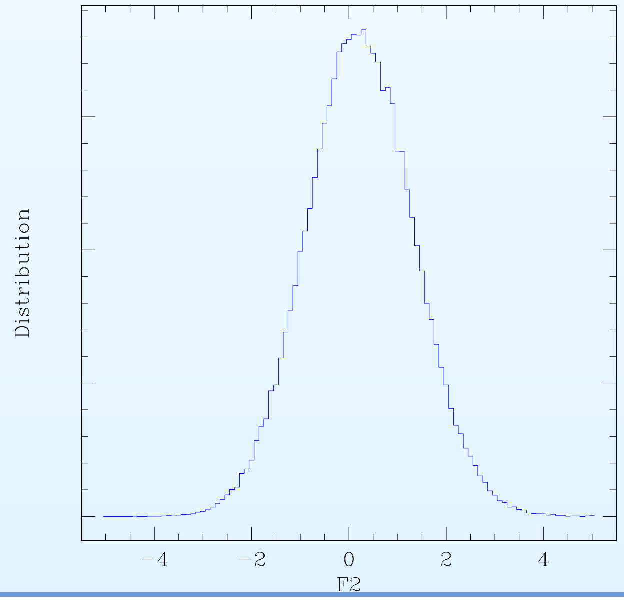
a posteriori test: there was no need to call the shell task (i.e. the orbital model is not worth keeping).

Single fit assessment

From the single star solution and its residuals (derived anyway):

- χ^2-test: $Pr(\chi^2 > \chi^2_\nu)$ under H_0: single star. $\nu = N - p$ where N is the number of observations and p the number of parameters.
- Goodness of fit:

$$F2 = \sqrt{\frac{9\nu}{2}}\left(\sqrt[3]{\frac{\chi^2}{\nu}} + \frac{2}{9\nu} - 1\right) \sim \mathcal{N}(0,1)$$



Both allow for direct comparison of fits with different models but the evaluation of $F2$ faster than of the probability.

Hipparcos single stars: $F2 \sim \mathcal{N}(0.22, 1.08)$

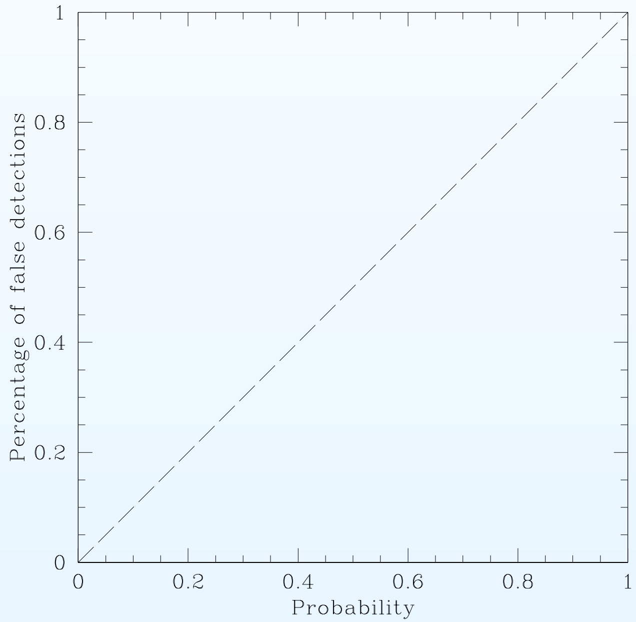
Improvement assessment

Whether it is called on purpose or *just in case*, the shell task offers an alternative model (with p' parameters). Accepting its results depends on how much they improve the fit.

Standard way of assessing the improvement:

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

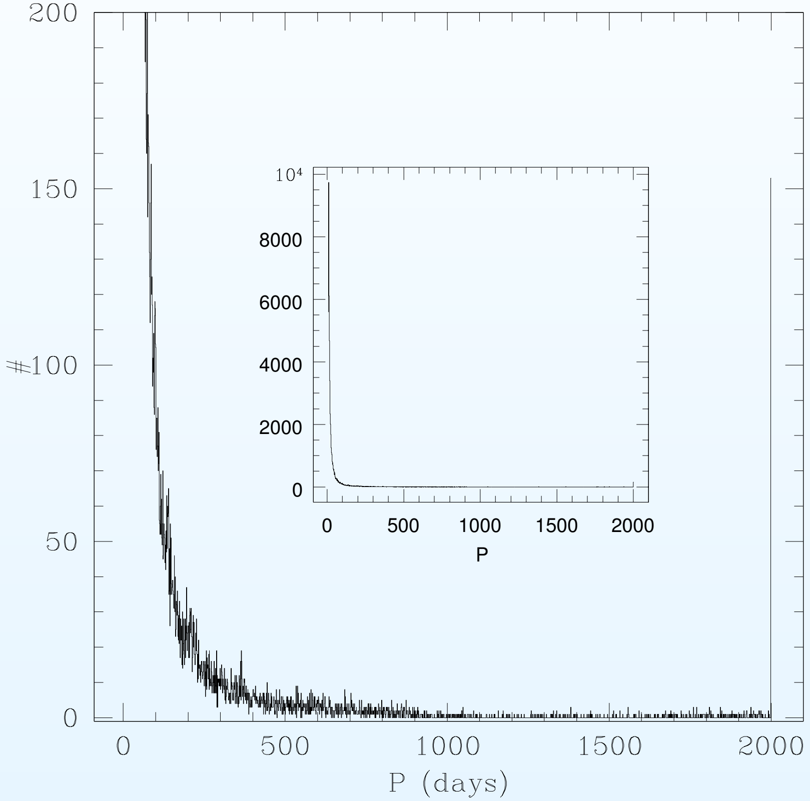
$$\alpha = Pr\left[\hat{F} < F(p'-p, N-p')|\text{no signal}\right]$$

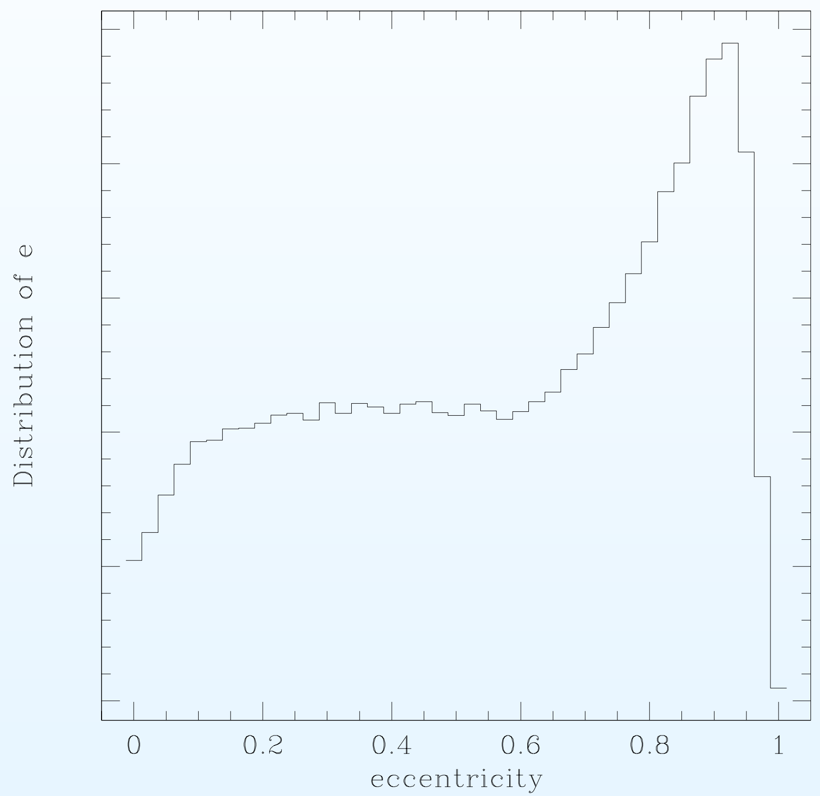


α sets the rate of false detection. Setting the threshold a bit high led to lots of pointless observations of some Hipparcos *binaries*. One billion stars with Gaia!

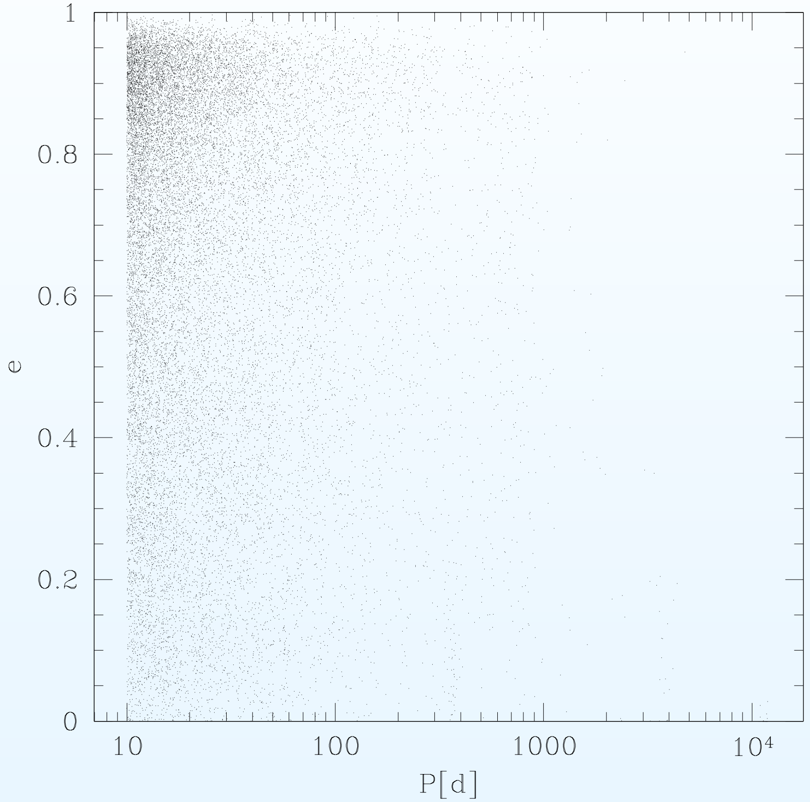
Just in case

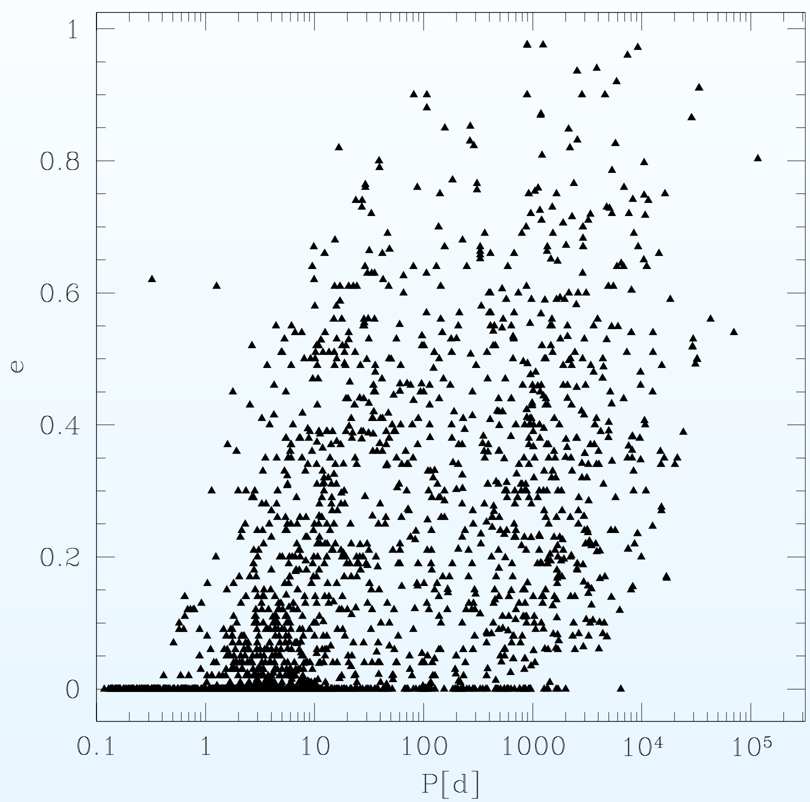
Orbital fit to the 100 038 *single* stars from the Hipparcos catalogue.





Awkward $e - \log P$ diagram

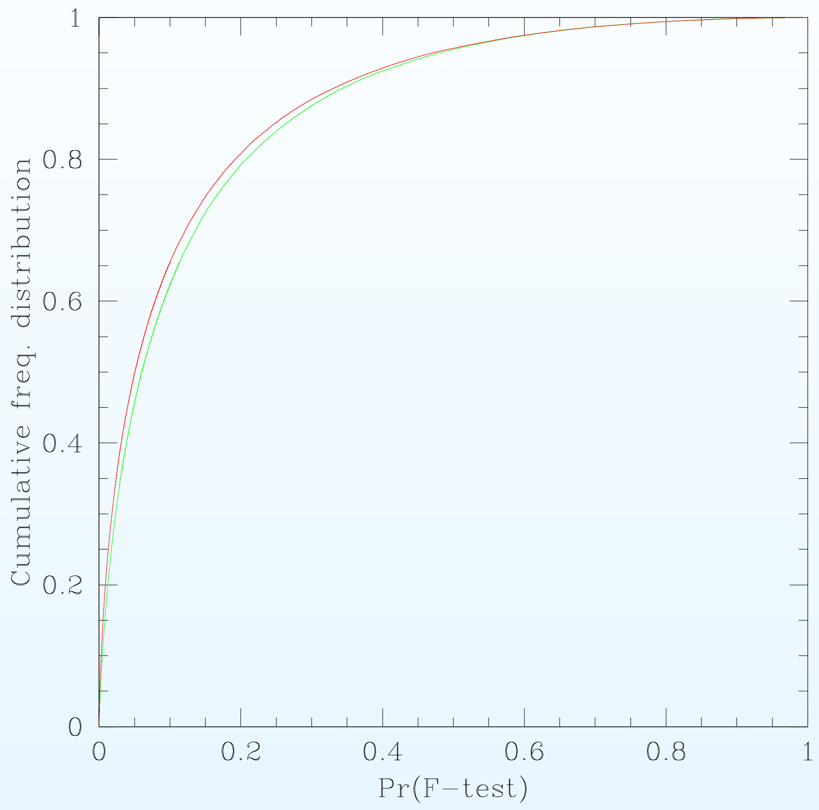


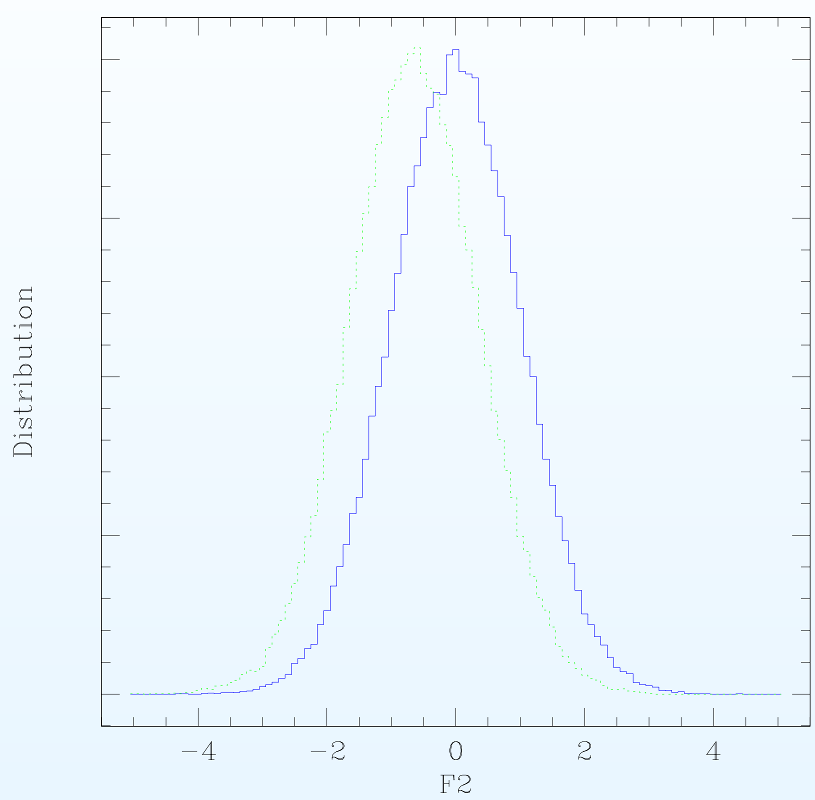


F-test @ 1%

NB: Orbits w with $P \sim 1$ year mimic the parallactic motion.

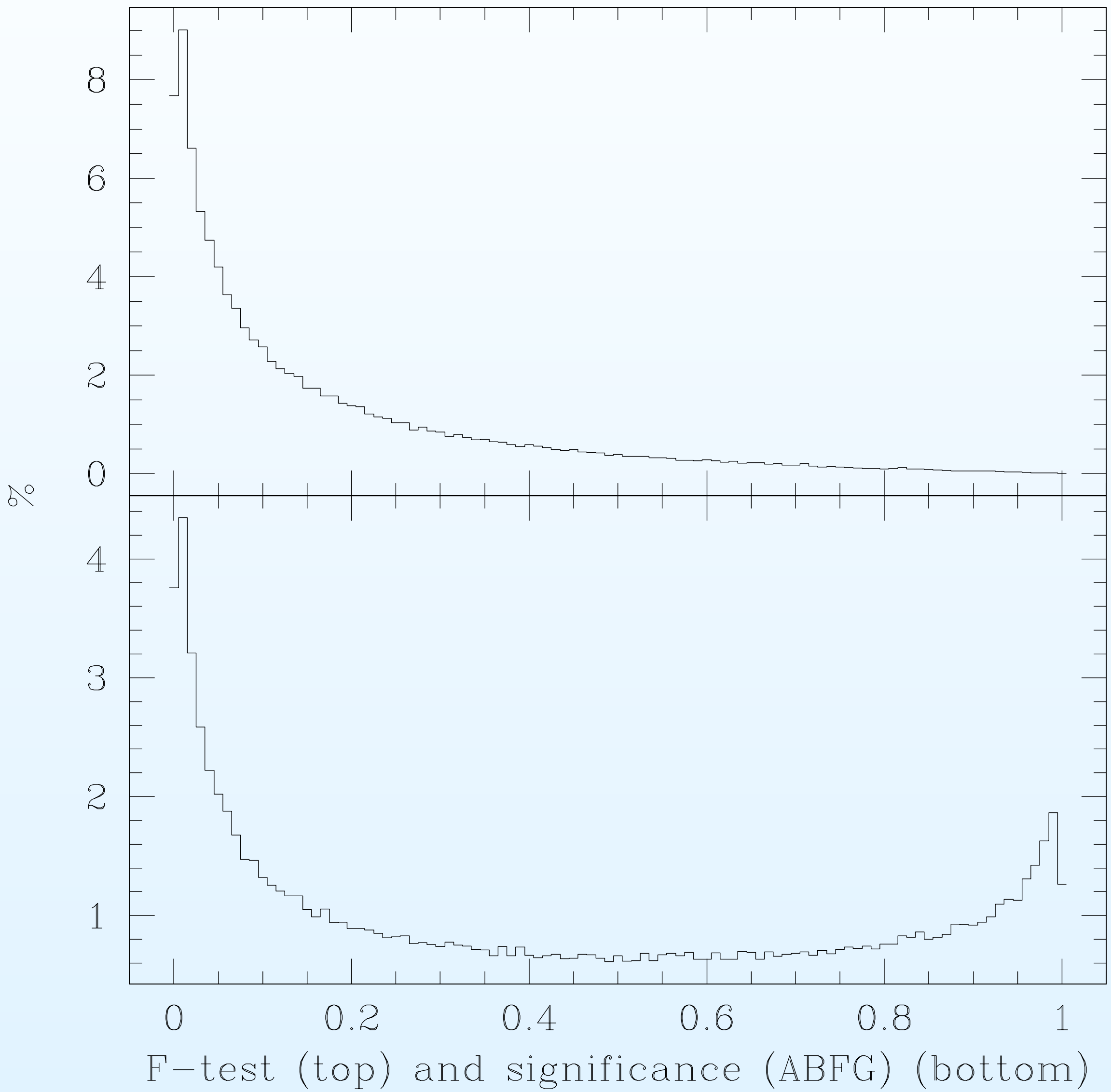
Statistical indicators





- 21 410 match the F-test @1%.
- 17 340 match the F-test @1% on synthetic data, $F2 \sim \mathcal{N}(-0.64, 0.996)$.

Synthetic single stars: 17k binaries (F-test 1%)



Reason of the excess

Let $\mathbf{v}$ denotes the observations, $\mathbf{a}$ the p parameters, $\mathbf{t}$ the N observation times and f the model:

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

is true iif f is linear.

Astrometric orbit model:

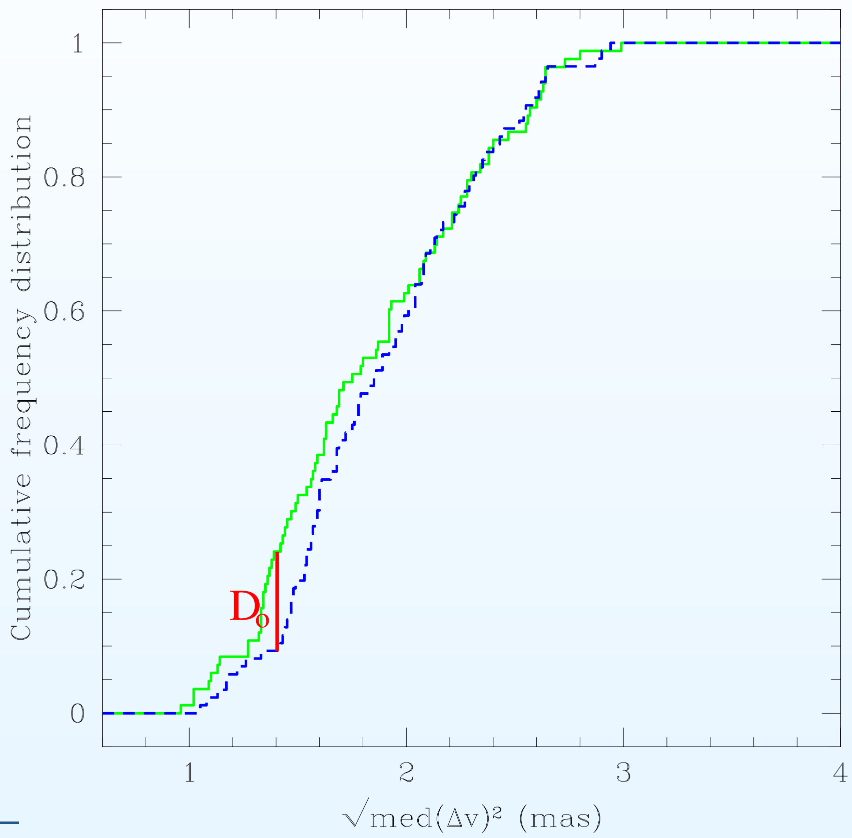
$$\begin{aligned}
x &= a_0(\cos\omega_1 \cos\Omega - \sin\omega_1 \sin\Omega \cos i)(\cos E - e) + \\
&\quad + a_0(-\sin\omega_1 \cos\Omega - \cos\omega_1 \sin\Omega \cos i)\sqrt{1-e^2}\sin E, \\
y &= a_0(\cos\omega_1 \sin\Omega + \sin\omega_1 \cos\Omega \cos i)(\cos E - e) \\
&\quad + a_0(-\sin\omega_1 \sin\Omega + \cos\omega_1 \cos\Omega \cos i)\sqrt{1-e^2}\sin E \\
\frac{2\pi}{P}(t - T_0) &= E - e\sin E
\end{aligned}$$

The model is not linear. $N-7$ therefore overestimates the number of degrees of freedom.

Kolmogorov-Smirnov test

Instead of comparing two values (e.g. χ^2), testing

- the observed distribution against a theoretical one,
- two empirical distributions against each other.



$$Q_{KS}(\lambda) \equiv 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2\lambda^2}$$

$$\begin{aligned}
N_e &= \frac{N_1 N_2}{N_1 + N_2} \\
D_o &= \max_{-\infty<x<+\infty} |S_1(x) - S_2(x)| \\
Pr(D > D_o) &= Q_{KS}(\left[\sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e}\right] D_o)
\end{aligned}$$

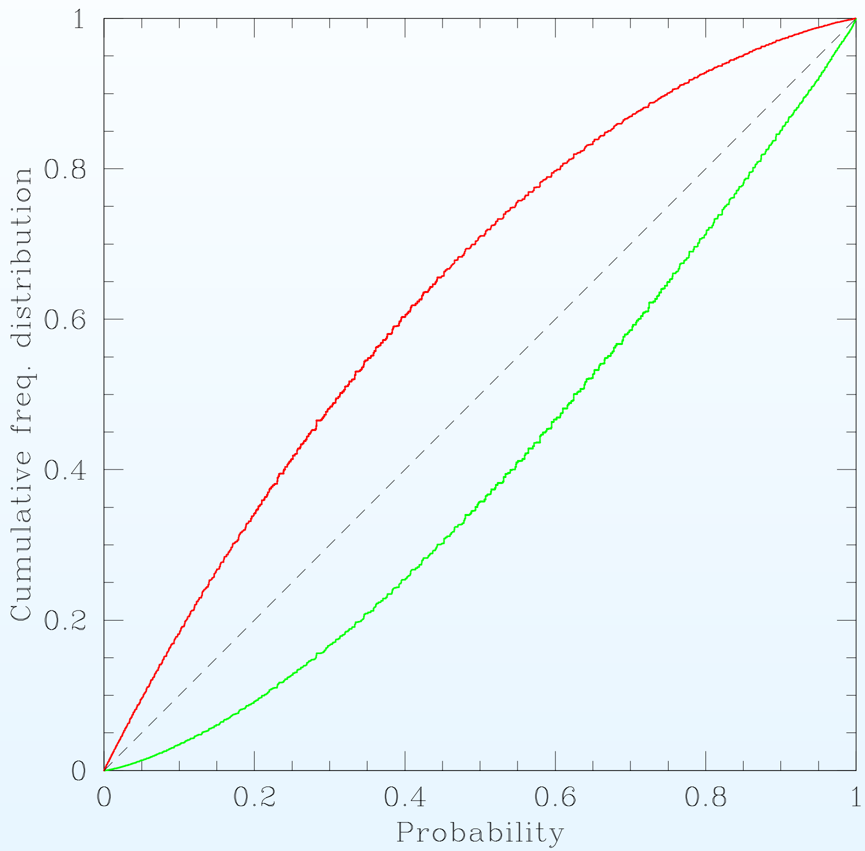
No more constraint on the distribution

KS test on two distinct distributions. Comparison of the

- weighted residuals (green),
- weighted residuals squared (red)

for the single star and orbital fits on synthetic single star data.

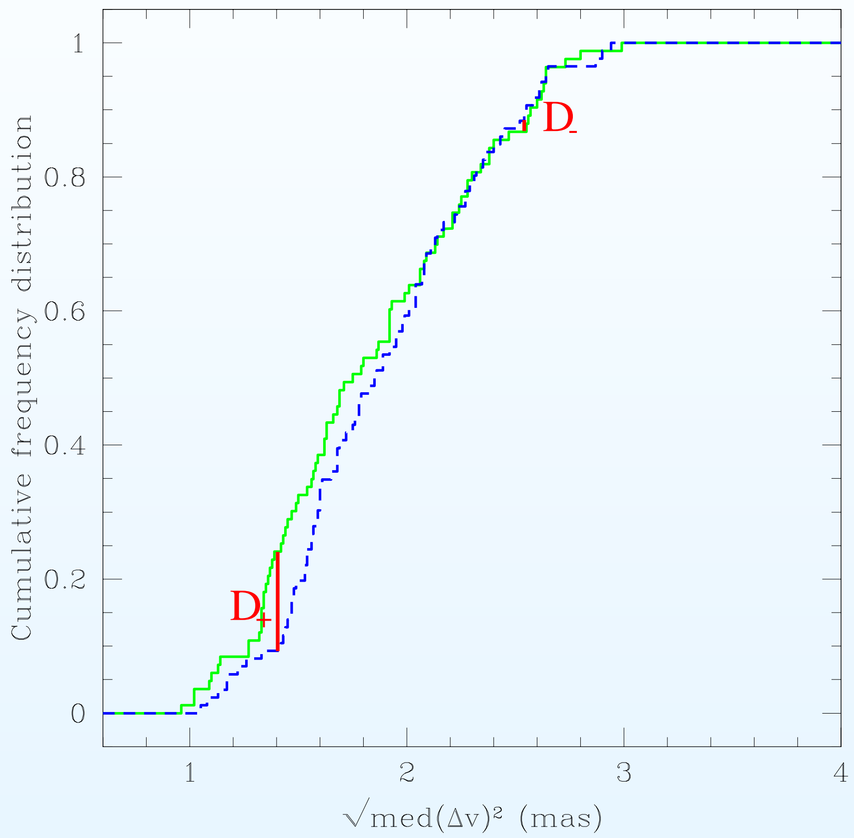
Maximum sensitivity of KS at the median.



- Orbital and single star residuals have same 0 median, so lack of highly significant discrepancies;
- Squaring makes the test over sensitive.

Kuiper test

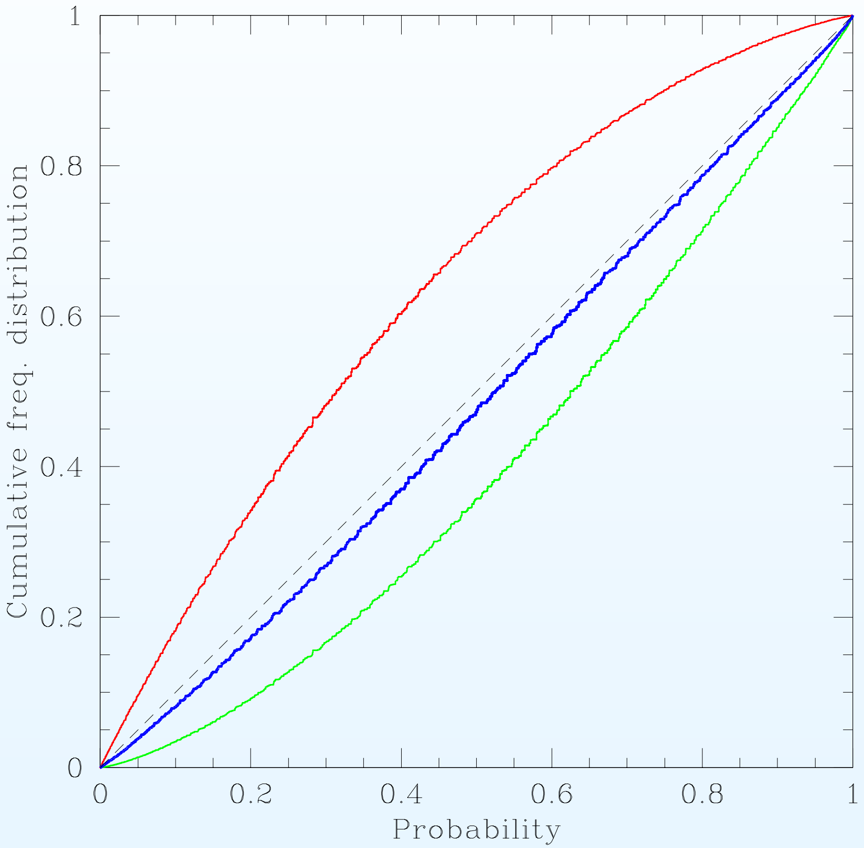
Kuiper test overcomes the sensitivity on the median of KS: no privileged point.



$$Q_{\mathrm{Ku}}(\lambda) \equiv 2\sum_{k=1}^{\infty}(4k^2\lambda^2 - 1)e^{-2k^2\lambda^2}$$

$$\begin{aligned}
N_e &= \frac{N_1 N_2}{N_1 + N_2} \\
V_o &= \max_{-\infty<x<+\infty}(S_1(x) - S_2(x)) + \max_{-\infty<x<+\infty}(S_2(x) - S_1(x)) \\
Pr(V > V_o) &= Q_{\mathrm{Ku}}(\left[\sqrt{N_e} + 0.155 + 0.24/\sqrt{N_e}\right] V_o)
\end{aligned}$$

Close to expectation



KS test on two distinct distributions. Comparison of the

- weighted residuals (green),
- weighted residuals squared (red)

and Kuiper test with weighted residuals (blue) for the single star and orbital fits on synthetic single star data.

Though Kuiper is a bit too severe, the overall behavior is much closer to the theoretical one. At 1%, 162, 2 054, and 721

F-test with a linearized model

Even if the model is nonlinear

$$\Xi^2 = (\mathbf{v} - f(\mathbf{a},\mathbf{t}))^t \mathbf{V}^{-1} (\mathbf{v} - f(\mathbf{a},\mathbf{t})),$$

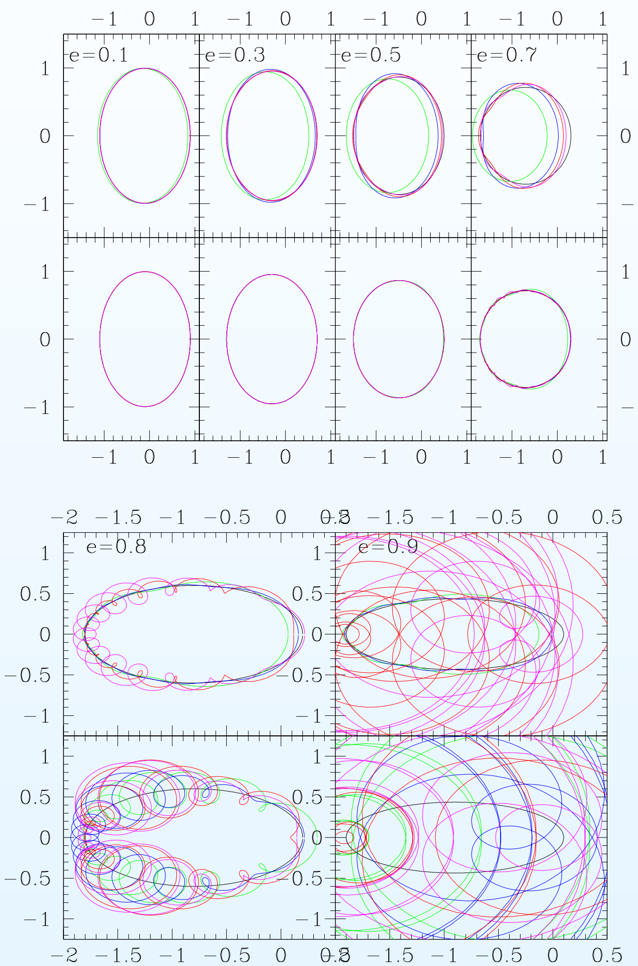
remains the *de facto* maximum likelihood estimator.

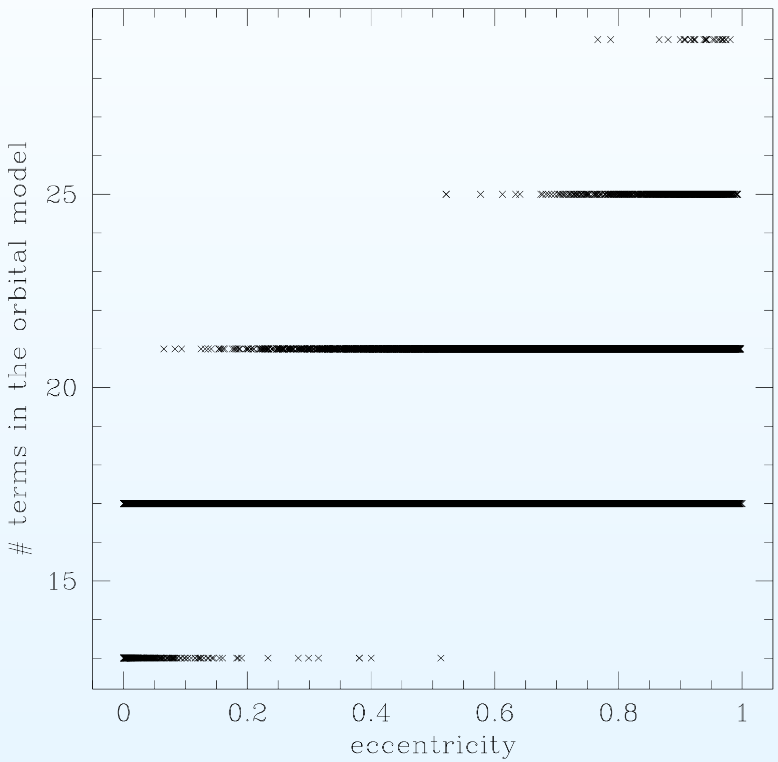
Substituting

$$\left\{ \begin{array}{l} X = \cos E - e \\ Y = \sqrt{1-e^2}\sin E \end{array} \right. \quad \text{with} \quad \left\{ \begin{array}{l} X = \sum_{k=0}^{\infty} c_{1,k}\cos(\frac{2\pi k t}{P}) + s_{1,k}\sin(\frac{2\pi k t}{P}) \\ Y = \sum_{k=0}^{\infty} c_{2,k}\cos(\frac{2\pi k t}{P}) + s_{2,k}\sin(\frac{2\pi k t}{P}) \end{array} \right.$$

makes the model linear in $c_{.,k}, s_{.,k}$ (A, B, F, G merge with those coefficients). The smallest linear model yielding the same Ξ^2 is adopted.

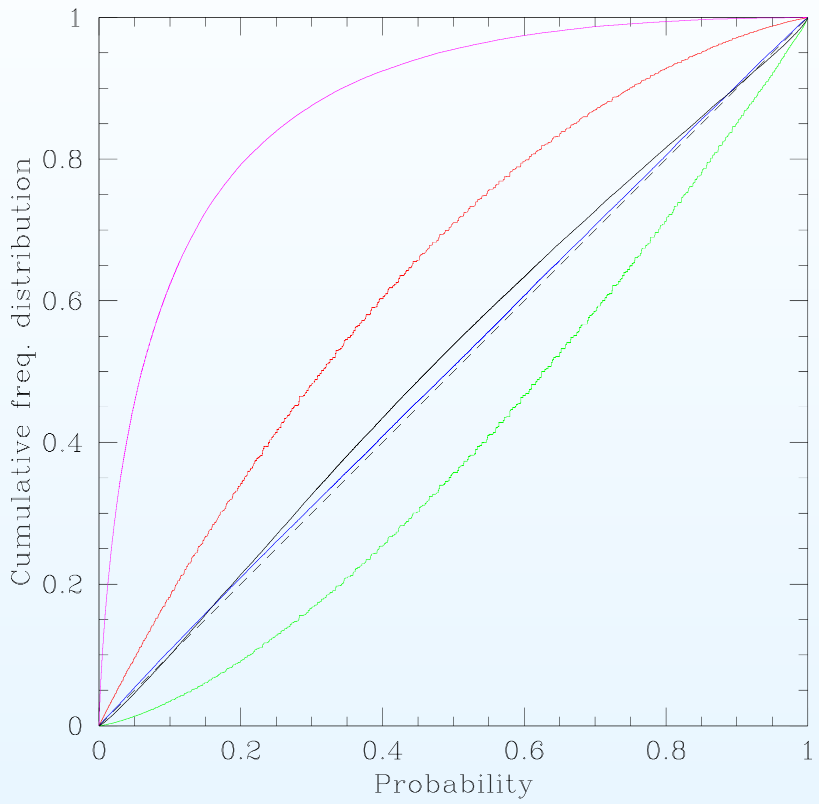
Number of degrees of freedom versus eccentricity

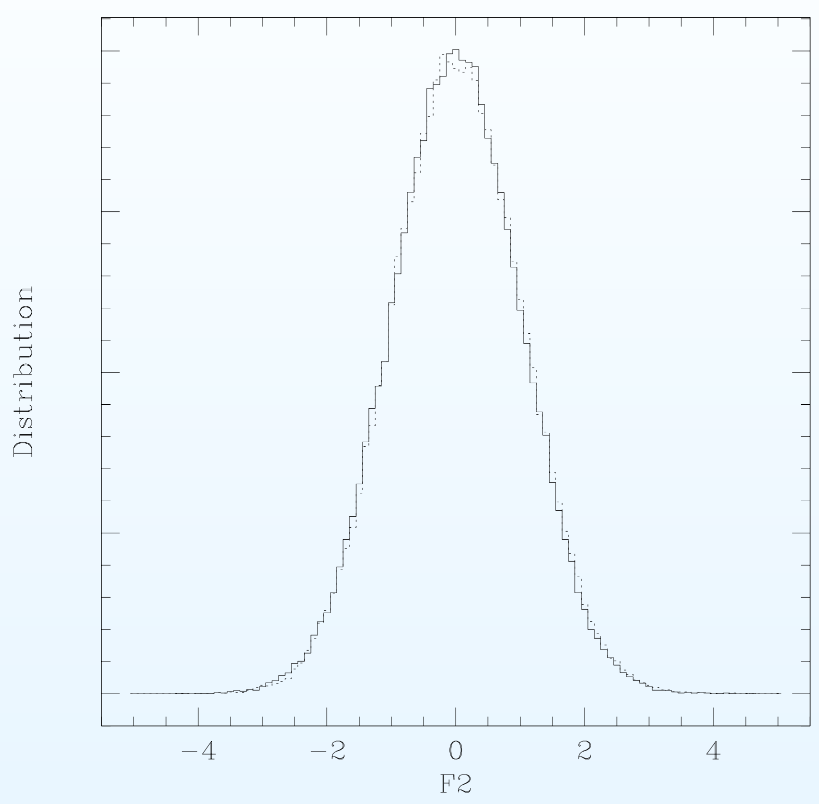




The number of terms generally increases with e but small data sets favor few terms in the Fourier expansion.

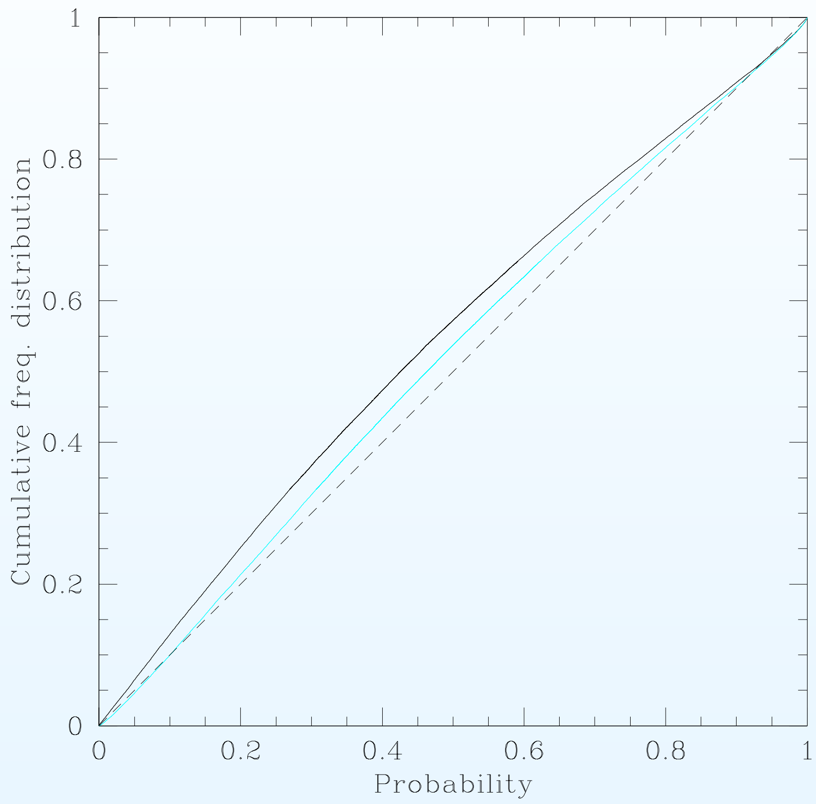
Results on synthetic data

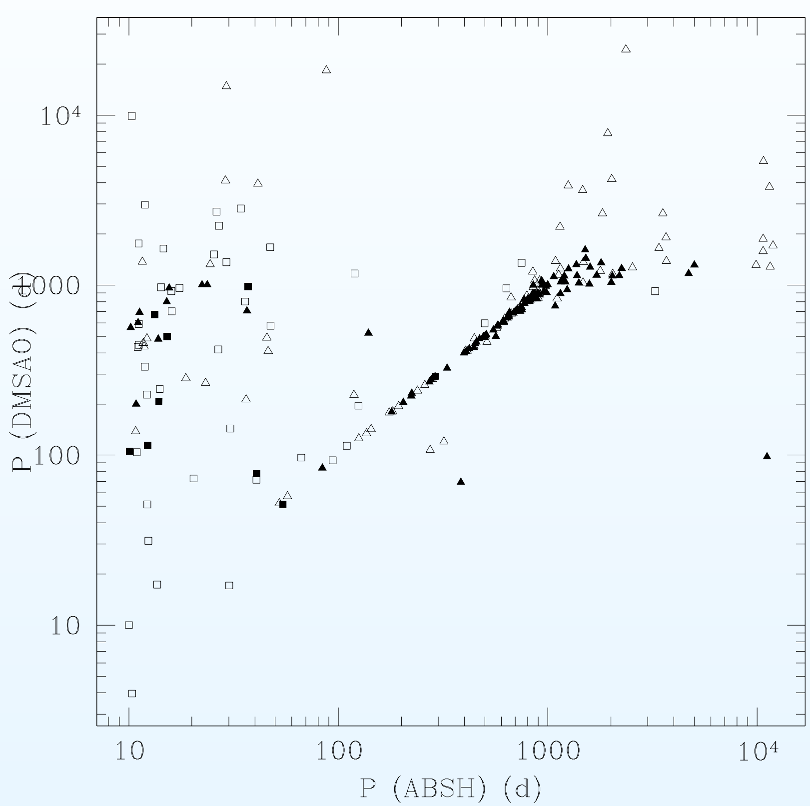




0.73% *binaries* with F-test @ 1% but 0.3% with $|F2| > 3$ ($F2 \sim \mathcal{N}(0.026, 1.00)$).

New and old binaries





With F-test @ 1%, 1.3% suspected binaries among the single stars and 76% binaries among the 235 DMSA/O. Some original DMSA/O solutions are questionable anyway (16% @ 1%).

Benefit from combining tests

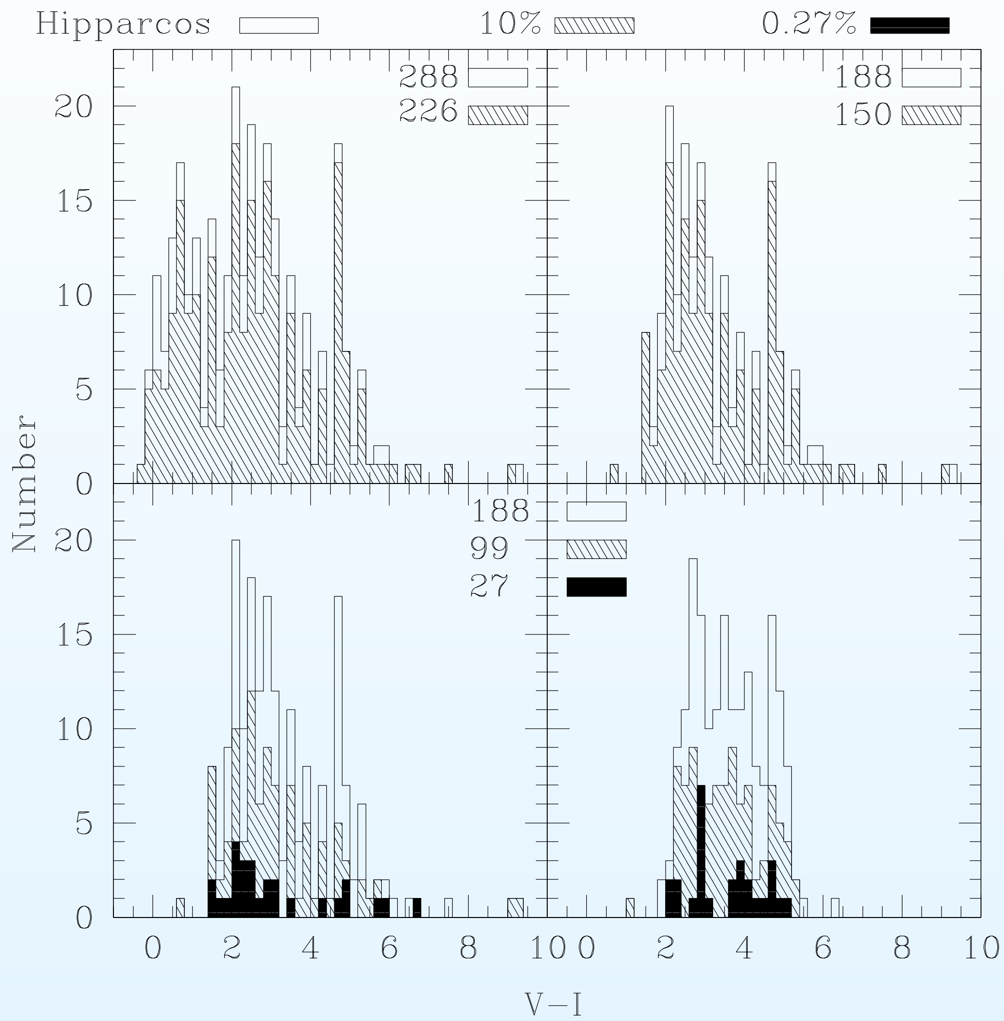
On synthetic data, F- and Kuiper tests poorly correlated: -0.08.

Results @ 1%:

Sample	Pr_S	F-test	Kuiper	KS+Kuiper
Synthetic	0.99	0.73	0.72	0
Real data	3.0	1.3	0.33	0.054
DMSA/O	84	76	74	67

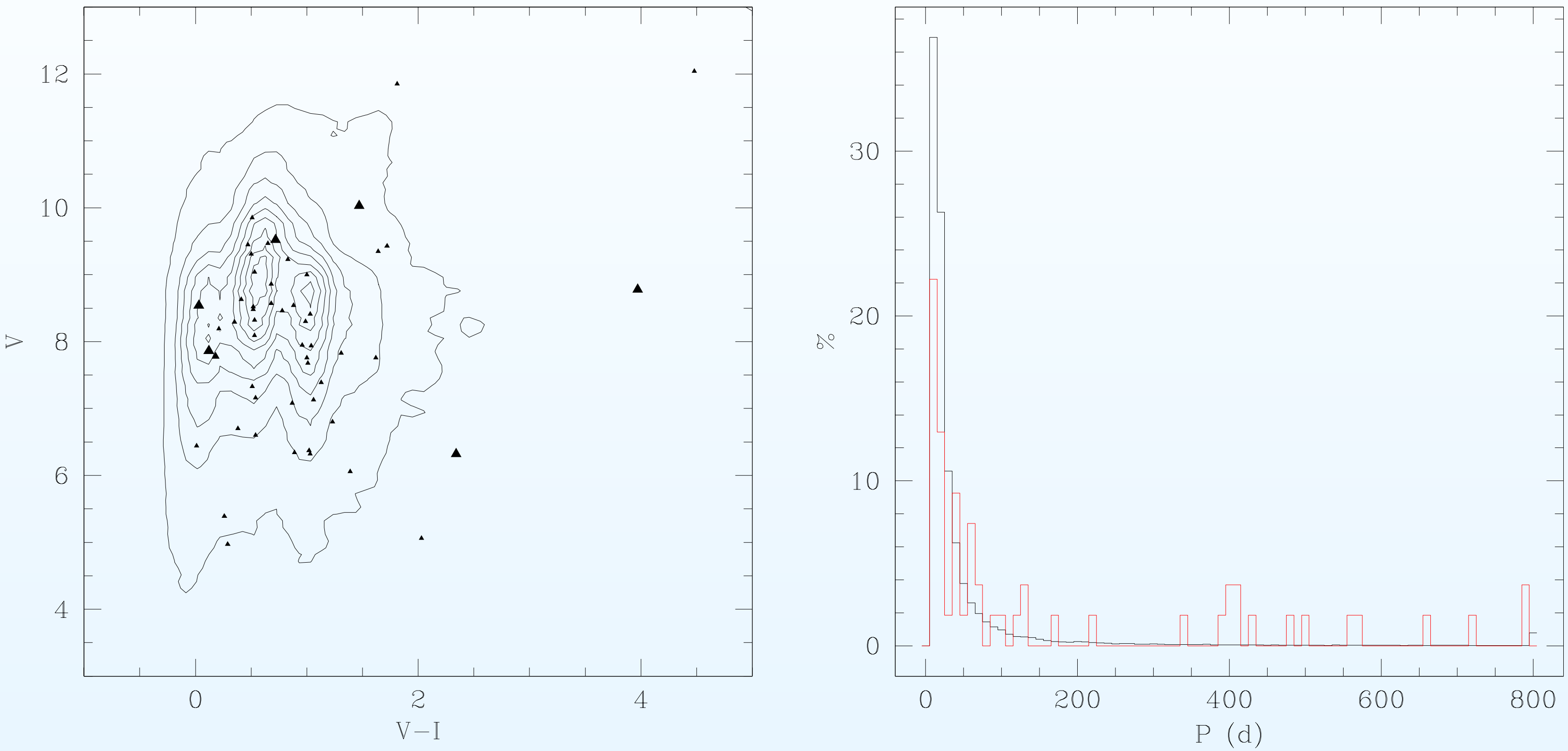
One expects ~ 10 false detections and, still, obtains 54 candidates.

Alternative explanation



Most Hipparcos VIM turned out to be single stars (Pourbaix et al. 2003)

Color and period distributions



Few candidates could be explained by weird colors and short periods.
Still, 14 are known binaries.

Conclusions

- Change of paradigm: it is no longer difficult to fit an orbit . . .
- . . . but figuring out whether it is worth keeping is.
- Murphy's law: a nonlinear model does not naturally turn into a linear one (if e is free, circular solutions are unlikely).
- The number of parameters in a nonlinear model underestimates the size of the equivalent linear model, thus jeopardizing the assessment of the fit.
- Theoretical investigations of the behavior of the statistical tests are needed (F-test might not be the only one which deserves revision).
- For the astrometric binary orbit (periodic phenomenon in general?), Fourier expansion offers a linear alternative to the nonlinear model. Simulations confirm the substitution is valid.
- Identifying the right linear model is affordable in terms of CPU time.
- 1% of 4M stars makes a lot of solutions to manually/visually check!