1 Purpose

The primary objective of the EPRV3 Evidence Challenge is to compare different algorithms and implementations for performing model comparison. We will provide 6 simulated radial velocity (RV) datasets, each with zero to three planets. Participants will be asked to compute quantitative estimates of Bayesian evidence (also known as the marginalized likelihood) for each model and the estimated uncertainty in each evidence, using whatever computational methods and simplifying assumptions that they choose. Participants are encouraged to apply multiple methods to analyze the same datasets and models.

The evidences for two models (and the same dataset) can be combined to compute the posterior odds ratio for any pair of n-planet to $(n+1)$-planet models. Any groups that are unable to estimate the evidence (e.g., using non-Bayesian methods), will be asked to provide estimates of quantities that they consider analogous to the posterior odds ratio. Of course, one should always provide an estimate of the uncertainties in the results that they report.)

We will compare results from different groups and different methods during the breakout sessions of the EPRV3 meeting in August 2017. We hope that this exercise will lead to a better understanding of the advantages and disadvantages of various algorithms, implementations and approaches. We are optimistic that this will lead to a paper describing the results, lessons learned, and recommendations for future studies.

2 Detailed Model Specification

In order to facilitate quantitative comparisons, it is important that each group use a standardized set of assumptions for the physical and statistical models. Therefore, we describe the process used to generate the datasets in detail. If you believe it is important to be provided with additional information, please let the organizers know as soon as possible.

2.1 Dataset Characteristics

We will provide 6 simulated datasets with anywhere from zero to three planets in each dataset. The datasets are generated with a set of consistent properties:

- Each dataset is a simple timeseries, including the time of observation (t_i), "measured" radial velocity (v_i), and a measurement uncertainty (σ_i).

- Number of observations: $n_{\mathrm{obs}} = 200$

- Observing baseline: 600 days

- Each dataset includes between zero and three planets (inclusive), as described below.

- Each dataset includes a single velocity offset and correlated, Gaussian noise to represent stellar activity.

2.2 Physical Model

In each dataset, the radial velocity of the star $(v(t))$ is computed via n-body integrations using Newtonian gravity, one star and zero to three planets. The simulated measurements are the sum of the line-of-sight velocity of the star and Gaussian noise. While the full model formally includes mutual planetary interactions, we fully expect that it would be well-described by the linear superposition of zero to three Keplerian orbits plus a constant velocity offset and a noise term. For the sake of computational efficiency, we restrict planets to orbital orbital periods ranging from 10 to 9,999 days.

The Gaussian noise is correlated from one observation to the next. The covariance matrix, Σ, is given by

$$\Sigma_{i,j} = K_{i,j} + \delta_{i,j} \left(\sigma_i^2 + \sigma_{Jitter}^2 \right), \tag{1}$$

where $K_{i,j}$ is a quasi-periodic kernel, $\delta_{i,j}$ is the Kronecker delta, σ_i^2 is the estimated measurement uncertainty for the ith observation, and σ_J^2 is the amplitude of an additional, unknown noise term. For the quasi-periodic kernel, we assume

$$K_{i,j} = \alpha^2 \exp \left[-\frac{1}{2} \left\{ \frac{\sin^2[\pi(t_i - t_j)/\tau]}{\lambda_p^2} + \frac{(t_i - t_j)^2}{\lambda_e^2} \right\} \right], \tag{2}$$

where the hyperparameters are fixed at the following values:

- $\alpha = \sqrt{3}$ m/s,
- $\lambda_e = 50.0$ days,
- $\lambda_p = 0.5$ (unitless), and
- $\tau = 20.0$ (days).

2.3 Statistical Model

2.3.1 Likelihood

Each simulated data point is generated according to

$$v_i = v_{\text{pred}}(t_i|\theta) + \epsilon_i, \tag{3}$$

where $v_{\text{pred}}(t_i|\theta)$ is the velocity predicted at time t_i by a model parameterized by θ, and ϵ_i is the perturbation to the measurement due to noise. $\epsilon \sim \text{Normal}(\mathbf{0}, \mathbf{\Sigma})$, i.e., the noise vector, ϵ is drawn from a multivariate process with covariance matrix Σ. Therefore, the appropriate likelihood $(\mathcal{L}(\theta) = p(\mathbf{v}|\theta))$ is a multi-variable normal distribution, centered on the predictions of the model (parameterized by θ),

$$\log \mathcal{L}(\theta) = -\frac{1}{2}(\mathbf{v} - \mathbf{v}_{\text{pred}}(\theta))^T \Sigma^{-1} (\mathbf{v} - \mathbf{v}_{\text{pred}}(\theta)) - \frac{1}{2}\log|\det\Sigma| - \frac{n_{\text{obs}}}{2}\log(2\pi), \tag{4}$$

where
    Therefore,

$$\mathbf{v} \sim \text{Normal}(\mathbf{v}_{\text{pred}}(\mathbf{t}), \Sigma), \tag{5}$$

where $\mathbf{v}$ is the vector of v_i's, and $\mathbf{v}(\mathbf{t})$ is the vector of true kinematic velocities at the vector of times $\mathbf{t}$. predicted velocities.

2.3.2 Priors

When computing the Bayesian evidence, we ask that each group adopt a common set of priors described below, so as to enable direct comparisons of their results. We assume a prior that factorizes in terms of each planet's orbital period (P_i), RV semi-amplitude (K_i), eccentricity (e_i), argument of pericenter (ω_i) and mean anomaly at epoch (M_i).

- For each planet's orbital period, $p(P) = \frac{1}{1+P/P0} * \frac{1}{\log(1+P_{max}/P_0)}$ where $P_0 = 1$ day and $P_{max} = 9,999$ days.

- For each planet's RV semi-amplitude, $p(K) = \frac{1}{1+K/K0} * \frac{1}{\log(1+K_{max}/K_0)}$ where $K_0 = 1$ m/s and $K_{max} = 999$ m/s.

- For each planet's eccentricity, $p(e) = 1$ from $0 \leq e < 1$

- For each planet's argument of pericenter, $p(\omega) = \frac{1}{2\pi}$ from $0 \leq \omega < 2\pi$ radians.

- For each planet's mean anomaly, $p(M) = \frac{1}{2\pi}$ from $0 \leq M < 2\pi$ radians.

- For the additional white-noise term, $p(\sigma_J) = \frac{1}{1+\sigma_J/\sigma_{J,0}} * \frac{1}{\log(1+\sigma_{J,max}/\sigma_{J,0})}$ where $\sigma_{J,0} = 1$ m/s and $\sigma_{J,max} = 99$ m/s.

- For the RV velocity offset, $p(C) \sim \text{Normal}(0 \text{ m/s}, 100 \text{ m/s})$.

3 File Formats

See the README.md in the data subdirectory for a description of the input data file format.

We will provide specifications for format of output files to facilitate rapid comparison of the results.