



# **BIP - Bayesian Inference with Python Documentation**

*Release 0.3*

**Flávio Codeço Coelho**

January 30, 2010



# CONTENTS

<b>1</b>	<b>Overview</b>	<b>3</b>
<b>2</b>	<b>Parameter Estimation in Dynamic Models</b>	<b>5</b>
2.1	Example Usage . . . . .	5
<b>3</b>	<b>Indices and tables</b>	<b>9</b>



Contents:



# OVERVIEW

The Bip Package is a collection of useful classes for basic Bayesian inference. Currently, its main goal is to be a tool for learning and exploration of Bayesian probabilistic calculations.

Currently it also includes subpackages for stochastic simulation tools which are not strictly related to Bayesian inference, but are currently being developed within BIP. One such package is the BIP.SDE which contains a parallelized solver for stochastic differential equations, an implementation of the Gillespie direct algorithm.

The Subpackage Bayes also offers a tool for parameter estimation of Deterministic and Stochastic Dynamical Models. This tool will be fully described briefly in a scientific paper currently submitted for publication.





# PARAMETER ESTIMATION IN DYNAMIC MODELS

A growing theme in mathematical modeling is uncertainty analysis. The Melding Module provides a Bayesian framework to analyze uncertainty in mathematical models. It includes tools that allow modellers to integrate Prior information about the model's parameters and variables into the model, in order to explore the full uncertainty associated with a model.

Once a model is thus parameterized, we can simulate the model, with full uncertainty representation and also fit the model to available data to reduce that uncertainty. Markov chain Monte Carlo algorithms are at the core of the framework, which requires a large number of simulations of the models in order to explore parameter space.

## 2.1 Example Usage

This first example includes a simple ODE (an epidemic model) model which is fitted against simulated data to which noise is added:

```
from BIP.Bayes.Melding import FitModel
from scipy.integrate import odeint
import scipy.stats as st
import numpy as np

beta = 1 #Transmission coefficient
tau = .2 #infectious period. FIXED
t0 = 0
tf = 36
y0 = [.999,0.001,0.0]
def model(*theta):
    beta = theta[0]
    def sir(y,t):
        '''ODE model'''
        S,I,R = y
        return [-beta*I*S, #dS/dt
                beta*I*S - tau*I, #dI/dt
                tau*I] #dR/dt
    y = odeint(sir,init, np.arange(t0,tf,1)) #np.arange(t0,tf,step)
    return y

F = FitModel(3000,1000,model,1,3,y0,tf,['beta'],['S','I','R'],
             wl=36,nw=1,verbose=False,burnin=1000)
F.set_priors(tdists=[st.norm],tpars=[(1.1,.2)],tlims=[(0.5,1.5)],
            pdists=[st.uniform]*3,ppars=[(0,.1),(0,.1),(.8,.2)],plims=[(0,1)]*3)
d = model(*[1.0]) #simulate some data
noise = st.norm(0,0.01).rvs(36)
dt = {'I':d[:,1]+noise} # add noise
```

```
F.run(dt, 'MCMC', likvar=1e-2, pool=True, monitor=['I'])
F.plot_results()
```

The code above starts by defining the models parameters and initial conditions, and a function which takes in the parameters runs the model and returns the output.

After that, we Instantiate our fitting Object:

```
F = FitModel(3000, 1000, model, 1, 3, y0, tf, ['beta'], ['S', 'I', 'R'],
             wl=36, nw=1, verbose=False, burnin=1000)
```

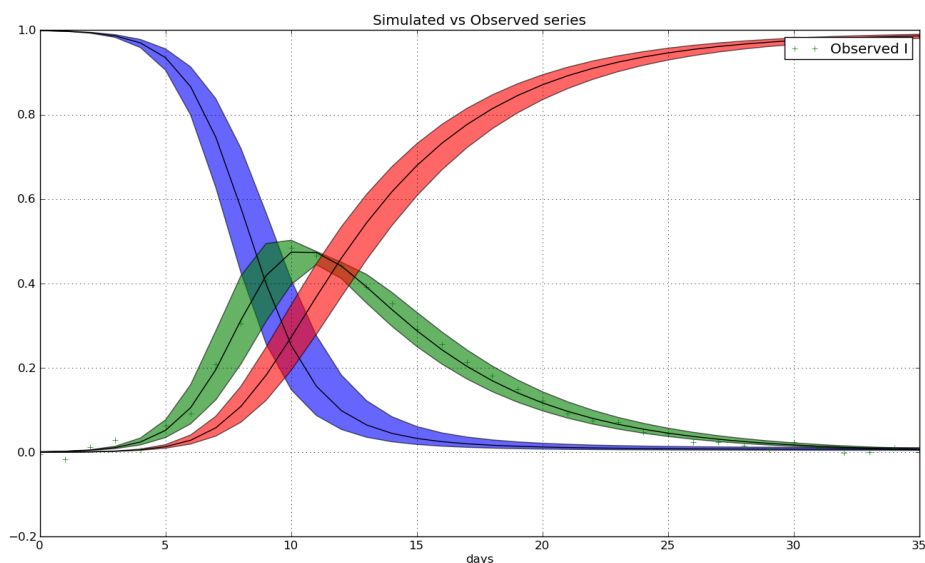
Here we have to pass a few arguments: the first ( $K=3000$ ) is the number of samples we will take from the joint prior distribution of the parameters to run the inference. The second one ( $L=1000$ ) is the number to retain if the inference method chosen (see below) in the sampling-importance-resampling (SIR) or approximate Bayes (ABC). then we have the number of parameters to be analyzed ( $n_{\theta}=1$ ), the number of output variables ( $n_{\phi}=3$ ), the initial condition vector ( $inits=y_0$ ), the list of parameter names ( $\theta_{names} = ['beta']$ ), the list of variable names ( $\phi_{names} = ['S', 'I', 'R']$ ), inference window length ( $wl=36$ ), number of juxtaposed windows ( $nw=1$ ), verbosity flag ( $verbose=False$ ) and finally the number of burnin samples ( $burnin=1000$ ), which is only needed for if the inference method chosen is MCMC.

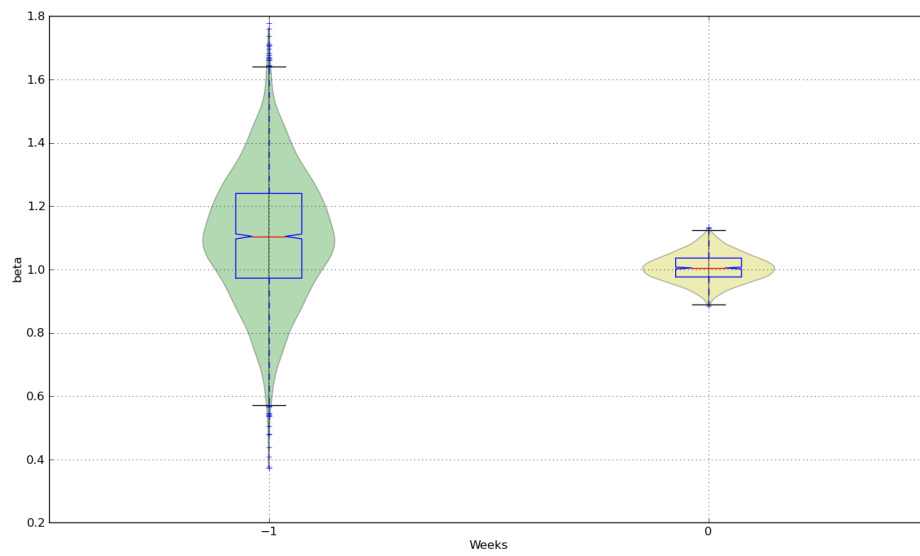
The next line of code also carries a lot of relevant information about the inference: the specification of the prior distributions. By now you must have noticed that not all parameters included in the model need to be included in the analysis. any number of them except for one can be set constant, which is what happens with the parameter  $\tau$  in this example:

```
F.set_priors(tdists=[st.norm], tpars=[(1.1, .2)], tlims=[(0.5, 1.5)],
            pdists=[st.uniform]*3, ppars=[(0, .1), (0, .1), (.8, .2)], plims=[(0, 1)]*3)
```

here we set the prior distributions for the  $\theta$  (the model's parameters) and  $\phi$  (the model's variables).  $tdists$ ,  $tpars$  and  $tlims$  are  $\theta$ 's distributions, parameters, and ranges. For example here we use a Normal distribution ( $st.norm$ ) for  $\beta$ , with mean and standard deviation equal to 1.1 and .2, respectively. we also set the range of  $\beta$  to be from 0.5 to 1.5. We do the same for  $\phi$ .

The remaining lines just generate some simulated data to fit the model with, run the inference and plot the results which should include plots like this:







# INDICES AND TABLES

- *Index*
- *Module Index*
- *Search Page*