Overview

This code performs a Bayesian analysis of decision making data. There are two metrics per decision: reaction time and choice accuracy (correct/incorrect). Further to this is the difficulty of the task and the degree of experimental interference with the decision making process. In our case, the difficulty is the log of the odour concentration ratio and the experimental interference is optogenetic stimulation (none, low, or high intensity light). There are also different genotypes representing the different control and experimental groups. The aim is to use the Bayesian Information Criterion (BIC) to elucidate firstly which decision making model best explains the evaluation of incoming information and secondly which parameter(s) within that model are most likely to be affected by the experimental manipulation of the decision making process. The two models that are compared are the extrema detection model (EXD) and the drift diffusion model (DDM). In the analysis presented in the publication (Wong *et al*), we determined that the DDM best explains the data and that the bound height parameter is dependent on the optical stimulation, while the drift rate and non-decision-time parameters are independent of it.

## **Installing and running the code**

This code was written and run using:

Windows 10

Intel Xeon processor (4 physical cores) @ 3.6 GHz

64GB Ram

Matlab 2019b, using the Statistics and Machine Learning, Distributed Computing, and Optimization toolboxes.

To install the software, copy all m files to a single folder and either set it as the Matlab working directory, or include it in the Matlab path.

To verify that the code works, a test data set is provided and there is a test mode that uses a smaller number of samples in the slicesample function. The test set contains the first 50 decisions from each genotype-optostim-difficulty level combination of the entire data set. Verifying the code in this way takes less than five minutes to run on the machine described above. To configure the code for test mode, open the file “script\_BayesSolve.m”, navigate to section 2 and change the value of “testingFlag” to true. Then run the script file from the Matlab command window (type “script\_BayesSolve” and press enter). A prompt window requesting the source data folder opens. Navigate to and select the parent data folder, “test set”. The script should run in a few minutes, printing one line of output for each genotype-stimulation-model combination. A folder in the working directory will be created called “output <time stamp>” that contains the posterior distributions of the various models for the drift diffusion model. There are also two tab-delimited text files summarising the output of both the drift diffusion and extrema detection modelling. The output should be identical to that in the folder “test output”.

To run the whole data set with the parameters used in the publication, set the “testingFlag” back to false and run the script. Select the data folder “behavioural parameters”. This took just under 17 hrs to run using the machine described above. The output should be identical to that in the folder “all data output”.

After running the code, relevant fitting parameters and the posterior pdfs will be saved to a file “fitting <time stamp>.mat”. This can be reloaded by following the instructions at the top of section 1 of script\_BayesSolve.

## Description of data structure

There are two metrics per decision: the reaction time and the choice (correct/incorrect). These are stored in the comma separated files: “reactiontimes.csv” and “trialaccuracy.csv”. There is one pair of files for each odour concentration ratio, optostimulation level and genotype. The directory structure is:

…\<genotype>\<optogenetic stimulation level>\<odour concentration ratio>\<file>

<genotype> is one of:

“NP6024 AMPAR (targeted Chrimson)”

“NP6024 CS (no Chrim)”

“NP6024 NLG1 (untargeted Chrimson)”

“VT030604 AMPAR (prime KC)”

The data from different genotype categories is analysed independently of each other (i.e. they are not grouped in any of the global models). The names of the folders have no value. If you wish to change, or add to, the genotype folders, update the variable “sources” in section 1 of the file “script\_BayesSolve.m” accordingly.

<optogenetic stimulation level> is one of:

'high intensity'

'low intensity'

'no light'

The optogentic stimulation is the basis for the experimental interference with the decision making process and so the data from the different stimulation levels are grouped together in the global models. These have no value (or order) and the stimulation folder names can be changed quite freely. However, the code is currently limited to accept a maximum of three different stimulation levels. To increase this, the function “globalHandle.m” will require modification. To analyse data from a single stimulation level, use only one optogenetic stimulation folder, and configure the script\_BayesSolve to run the independent model only (in section 2, change globalFitInds to [] and keep fitIndepFlagset to true).

<odour concentration ratio> is in the format ‘<LHS flow rate>vs<RHS flow rate>’ where the flow rates take the values 10, 50, 70, 90 or 100.

The difficulty of the task is calculated in the script as logDiff = |log(<LHS>/<RHS>)| in section 1 of script\_BayesSolve. If you wish to change this, modify the parameter logDiff.

## Description of global models

tND and sigmaND refer to the mean and standard deviation of the non-decision time. A is the bound height, or evidence threshold. k is the drift rate in the DDM, or the standard deviation in the evidence of successive snapshots in the EXD. The independent models are those that are run separately for each optical stimulation level and so have separate A, k, tND and sigmaND for each stimulation level. The global models assume that some or all parameters are the same across the different optostim levels as summarised in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model number | A | K | tND | sigmaND |
| 1 | Free | Free | Global | Global |
| 2 | Global | Free | Global | Global |
| 3 | Free | Global | Global | Global |
| 4 | Global | Global | Global | Global |

It is possible to modify the decision making models by changing the functions “pDDM” or “logP\_RT\_acc” and the “globalHandle” function.

## Description of Output

After running the script, a folder is created in the working directory called “output <date stamp>” where date stamp is in the format “yy-mm-dd HHMM”. This folder contains the posterior pdfs of the DDM models. The files are named “DDM M<X>. <text description>. <genotype> <optostim level>”. The model <X> is numbered 0 for independent, or 1 to 4 for the global models as listed in the table above. The text description lists the global and free (individual) parameters. The genotype is determined as described in the description of data structure. For model number 0, the optostim level of the input data set is appended. In addition to the posterior pdfs are summary output files containing the means and confidence intervals for the fitted parameters and the BIC for each model. There is a summary file for the DDM and a summary file for the EXD. The text file is in tab delimited format which can be neatly copied into your favourite spreadsheet package.