CCN Assignment 2
Computational modelling of behavioural data

Filippo Ferrari s1556895@ed.ac.uk
Based on work by Sam Rupprechter

Spring 2021

Practicalities

Deadline is Monday, **March 22 2021** at **4 pm** (standard late policies apply). Please submit your report as a **single PDF** on **Learn**.

- Please name your submission as `s1234567_ccn_cw2.pdf` where `s1234567` is your student number. The submission should be no longer than **8 pages**. References, appendices and code do not count toward this limit.
- You do have to **include code in your report**. Please attach it to your PDF as an appendix in the same way you did for Assignment 1.
- There is a bonus task which you can safely ignore. Only very few extra marks are available but you might learn a useful technique for computational modelling. Your total marks can not go above 100.
- Plots should always include axes labels and units. Figures should always have a caption and be referenced in the text. The presentation and format will count towards the final mark. Note that using default settings for plots is probably not the best idea (e.g. the line-widths are often too thin).
- Show your work. If there are more than one (sensible) ways to calculate something be sure to describe how you did it. Be concise and precise in how you report your results. Don't include lots and lots of separate graphs: you can superimpose different graphs in the same plot.
- Copying results is not allowed. It is okay to ask for help from your friends. However, this help must not extend to copying code or written text that your friend has written, or that you and your friend have written together. You are assessed you on the basis of what you are able to do by yourself. If you are found to have done so, a penalty will be assessed against you as well.
- Your programming style will not be assessed but your code should be readable/decipherable in a reasonable amount of time. If your code is incorrect, you may still be able to get points by realising that the results are not correct and describing how you think they should look like.
- The two parts are independent and can be completed in any order. They focus on different types of models and techniques used in computational psychiatry. Within each part it is recommended that you try to complete the tasks in the order they are given. Please clearly indicate which question you are answering in your report.
- If something is not clear please create a new post on Piazza so that the answer can be shared with all the other students. If unsure whether your question can be shared with all the students, you can create a private Piazza question.

Part 1: Drift Diffusion Process (23 points)

Depression is associated with numerous cognitive deficits, including deficits in executive functioning, memory and attention (Rock et al., 2014). Additionally, we have learned that one of the symptoms of a major depressive episode is "psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down)" (according to DMS-V; Edition et al. (2013)). Sometimes we therefore observe worse performance in depressed patients as compared to healthy control participants (e.g. in the tasks of the second and third parts of this assignment). Other times, we observe similar performance in terms of correctness, but slower response times from depressed participants.

We can model the dynamics of decision processes using a random walk model or a diffusion process (see lectures slides, and Smith and Ratcliff (2004) for a review). This model starts from the basic principle of accumulation of information. When an individual is asked to make a binary choice on the basis of an available stimulus, the assumption is that evidence from the stimulus is accumulated over time and a decision is made as soon as an upper or lower boundary is reached. Which boundary is reached determines which response is given and the time required to reach it determines the response time (RT).

In this part, we will use a drift diffusion model to model responses of depressed and healthy participants in some (for our purposes arbitrary) experiment.

Introduction

We model a decision process between two hypotheses which we call h^- and h^+. A Wiener diffusion process W_t with drift rate v and variance s^2 can be simulated using the following equation (Ratcliff and Rouder, 1998):

$$W(t + dt) = W(t) + v \times dt + s \times \eta \tag{1}$$

When $W(t) < 0$, a decision in favour of h^- is made. On the contrary if $W(t) > a$, a decision is favour of h^+ is made. dt denotes the time step used to simulate the process. η represents a Gaussian noise process with standard deviation $\sqrt{dt} : \eta \sim \mathcal{N}(0, dt)$. (Note the square root!) The parameters are: the mean drift rate v, the separation of the boundaries a, and the starting point $W(0) = z$. The drift rate v models the amount of evidence in favour of each of the hypotheses. If $v > 0$, there is more evidence in favour of h^+ (which is then the correct decision). On the contrary, if $v < 0$, there is more evidence in favour of h^-.

Task (a): Initial Simulations (9)

Simulate the model with the following values: $v = -0.04$, $a = 0.15$, $s = 0.04$, $z = a/2$, $dt = 0.001$. Implement a "timeout" of a maximum of 2000 steps (where one step corresponds to one W update). (Repeat the simulation multiple times.)

- Plot example paths.
- Plot the three response time distributions (h^+, h^- and overall histograms).
- Report and discuss the percentage of correct and incorrect responses and timeouts. Can you infer something about the experiment?

Task (b): Exploring parameter settings (5)

Explore different settings of these values for v and a (always set $z = a/2$). Illustrate and describe how changes in parameter settings influence accuracy, response times and their trade-off.

Task (c): Prior information (4)

Assume that a subject knows that one hypothesis is more likely than the other, for example the prior $p(h_+) = 3 \times p(h_-)$. How can this information be included in the model? Illustrate and explain.

Task (d): Group differences (5)

Assume we collected data from two different experiments (with different participant groups). Each experiment included a participant group suffering from major depressive disorder and a healthy control group. Discuss the following scenarios.

- Using two separate statistical tests to compare groups in terms of accuracy and response time, we did not find a significant difference for either test. We then fitted a drift diffusion model and compared the estimated parameters of the two groups. Is it possible that we would find significant differences between the groups' parameters and what would it mean for the sensitivity of the modelling analysis? How do DDMs compare to reinforcement learning models in this regard?
- We find that (compared to controls) patients responded slower during the task but with similar accuracy. We suspect that this is related to one of two things: (a) A difficulty in the evidence accumulation and/or integration due to some cognitive deficits, or (b) a "slowing down" of their perception and/or movement initiation and execution. How could we test for these two possibilities using our model and/or an extension of the model?

Part 2: Model Fitting (77 points)

As you have learned in the lectures, there is evidence that reinforcement learning is impaired in patients suffering from major depressive disorder (MDD; see Chen et al. (2015) for a review). Important neural evidence comes from functional magnetic resonance imaging (fMRI) studies, which have shown both reduced prediction error signals as well as expected value signals in the striatum and other areas of the brain (Kumar et al., 2008; Gradin et al., 2011). In this assignment we will focus on modelling *behavioural* impairments during a probabilistic reversal learning task inspired by an article of Dombrovski et al. (2010). In their experiment, participants had to choose between two coloured rectangles on each of the 80 trials in order to maximize their reward. Each choice was followed by a reinforcer (a symbolic reward or punishment in the form of a green or red frame around the stimulus and a high- or low-frequency tone), from which participants could learn. One of the stimuli was followed by reward in 80% of trials (and followed by punishment in the remaining trials), while the alternative choice led to punishment 80% of the time. These contingencies reversed after 40 trials so that participants had to re-learn them. Dombrovski et al. (2010) fitted a reinforcement learning model to the behaviour of participants from multiple groups (depressed suicide attempters, depressed suicide ideators, non-suicidal depressed patients, healthy controls) and found that groups differed in one of the fitted parameters of the model ("memory" parameter). For this assignment, we will focus on a simpler version of the experiment, in which no punishments were presented (i.e. if the outcome was not a reward, it was just an empty screen). We will also implement slightly different reinforcement learning models.

Introduction

You are given data from a simple reversal learning experiment, which you can download from the CCN Learn webpage. (Note that this data was simulated.) The experiment was performed by 24 MDD patients and 31 healthy control participants matched for age, sex and IQ. In each trial of the experiment, participants were asked to make a simple choice between two stimuli. Each stimulus had a certain probability of being followed by a reward (as opposed to no reward). The probabilities were unknown to participants and changed every 24 trials, switching between pairs of probabilities. The participants were unaware of these switches. The first pair had a reward probability of 40% (for stimulus A) and 85% (for stimulus B). The second pair had a reward probability of 65% (for stimulus A) and 30% (for stimulus B). After that it switched again to the first pair and then the second pair and so on. Note that stimulus A in both pairs always appeared the same to the participants. Only the reward probability changed and participants were not informed about that. They simply had to choose one of the two stimuli on each of the trials. Participants were asked to maximize their reward and each participant completed 240 trials.

We will model the behaviour of participants using a simple reinforcement learning model, in which the value of the chosen stimulus i will be updated on trial t after observing reward r (which will be 0 for no reward and 1 for reward) as follows:

$$V_i^{(t+1)} = V_i^{(t)} + \varepsilon \times (r^{(t)} - V_i^{(t)}) \tag{2}$$

Initial values $V^{(0)}$ are set to zero. When the outcome of one stimulus choice is observed, the value of the alternative option is not updated. The probability of choosing stimulus A as opposed to stimulus B on trial t is modelled using a softmax function (written slightly differently than in the lectures):

$$p(action\ A \mid V^{(t)}, \beta) = \frac{\exp(\beta \times V_A^{(t)})}{\exp(\beta \times V_A^{(t)}) + \exp(\beta \times V_B^{(t)})}. \tag{3}$$

We therefore have two parameters: the learning rate ε in our observation model and the inverse temperature β in our decision model. We hypothesise that depression could be related to changes in reward learning (i.e. an altered learning rate) or some change in decision making (inverse temperature).

Task (a): Exploring the data (4)

Load and explore the provided data: For each participant, calculate the number of received rewards and the number of times they chose stimulus B. Did our participants perform the task well? What is the expected number of received rewards for participants who respond randomly throughout the experiment?

Task (b): Simulations (7)

Since we are working with a generative model, we are able to simulate data, which can be extremely useful. Use equations 2 and 3 to write a function that lets you generate data from known parameters (the parameter values should be an input to the function). Generate the rewards with probabilities which are changing every 24 trials as described in the introduction.

Hint: You may use the code from the lectures to help you get started.

Simulate 240 choices with parameter settings $\varepsilon = 0.35$ and $\beta = 5.5$ a number of times. (Choose a reasonable number so you can average the simulations). Illustrate the average evolution of values $V(A)$ and $V(B)$. Illustrate the average evolution of the difference in V values of the two stimuli (i.e. show how $V(A) - V(B)$ changes, on average, over the course of the simulated experiments). Very briefly explain what is observed and why the shape of this evolution makes sense.

Hint: The average number of received rewards from a large number of simulations using the parameter values stated above should be around 145–155.

Task (c): Exploring parameter settings (6)

Simulate 240 choices several times for a number of different parameter settings. Systematically vary settings of ε and β to explore how different values affect the average number of simulated received rewards. Plot the average number of rewards as a function of the parameter settings, which means there will be three dimensions. Choose sensible ranges for the parameters (e.g. $0 < \varepsilon < 1$ and $1 < \beta < 15$).

Briefly describe your precise approach and comment on how the expected performance during the experiment is related to different settings of the parameters.

Task (d): Likelihood function (6)

To find the parameter values that best capture each participant's behaviour, we need to define the likelihood of the parameters. Write a function that takes as input the data (choices and rewards) for an individual and a vector of parameters (learning rate and inverse temperature) and returns the *negative log likelihood (NLL)* of these parameters (note that θ is the parameter vector containing both ε and β):

$$NLL = - \sum_{c \in Choices} \log p(c \mid V, \theta). \tag{4}$$

Hint: You may use the code from the lectures to help you get started.

Report the NLL for the second (patient) participant using parameter settings $\varepsilon = 0.4$ and $\beta = 6$.

Hint: If you compute the NLL for the first (patient) participant using $\varepsilon = 0.4$ and $\beta = 6$, it should be close to 132.

Task (e): Model fitting (7)

Find the parameters that minimize the *NLL* for each individual: Pass your NLL function and a set of starting parameters (use $\varepsilon = 0.5$ and $\beta = 5$) to a gradient-based optimization function (performing unconstrained minimization).

Hint: For MATLAB, use `fminunc`. It will automatically approximate the gradient using finite differences. For Python, you could have a look at `scipy.optimize` and particularly `fmin_bfgs`.

Calculate and report mean and variance of the fitted parameter values for learning rate and inverse temperature. Illustrate the results: Plot the participant index on the x-axis, parameter values on the y-axis. Make sure the different groups are easily distinguishable. Comment briefly on what you observe. Do all the values make sense? If not, what can you do about it? Calculate and report the Pearson's correlation coefficient between estimated parameters across all participants. Calculate and report the Pearson's correlation coefficient between estimated parameters separately for participants within each group.

Task (f): Group comparison (5)

Use a two sample t-test (describe which exact test you are using) to test whether the estimated parameter values (Task (e)) are significantly different across groups. Report the t statistic, degrees of freedom and p value if applicable. Make sure the estimated parameter values you use for this test make sense. Explain briefly how we can interpret the results and how this relates to our hypotheses. What might we conclude if the data was real?

Task (g): Parameter recovery (8)

We now want to check the reliability and identifiability of our parameter estimates. Sample 55 sets of parameter values of learning rate and inverse temperature from a multivariate normal distributions. Choose sensible numbers for the mean of this distribution and describe how you chose them; choose small numbers for the variance (e.g. 0.01 and 0.5 for ε and β respectively); set the covariance to zero. Illustrate the sampled values and highlight and exclude (resample) nonsensical values.

Use the sampled parameter values to simulate 55 sets of data (as in Task (b)). Fit new parameter values to these simulated data sets (as in Task (e)). Calculate, report and illustrate the Pearson's correlation between the parameter values you used to simulate the data and the parameter values that you obtained from fitting the model to the simulated data.

Comment briefly. Does this parameter recovery simulation meet your expectations? Explore and describe how the number of trials and the number of simulated data sets affects the performance of the parameter recovery.

Task (h): Alternative models (8)

In addition to *model 1*, which you already implemented, we will now be considering two additional models.

Model 2 replaces the inverse temperature parameter β by a *reward sensitivity* parameter ρ so that we obtain the following two equations:

$$V_i^{(t+1)} = V_i^{(t)} + \varepsilon \times (\rho \times r^{(t)} - V_i^{(t)}) \tag{5}$$

$$p(action\ A \mid V^{(t)}) = \frac{\exp(V_A^{(t)})}{\exp(V_A^{(t)}) + \exp(V_B^{(t)})}. \tag{6}$$

Model 3 has the same equations as model 1, but adds two additional parameters representing the initial values v_A and v_B. (Set $V^{(0)} = [v_A, v_B]$ instead of setting $V^{(0)} = [0, 0]$).

1. Implement simulation and (negative) log likelihood functions for each of these two new models. (Copy the functions you wrote for model 1 and make the required changes, which should be few.)
2. Fit the two new models. Illustrate the fitted parameter values. Are the results surprising or as expected? Can you explain them?

Task (i): Model comparison (6)

For each participant you should now have three negative log likelihood values: One for each model (using the estimated optimal parameters). Compare these negative log likelihood values between models. Explain what you observe. Does it make sense?

For each participant, compute AIC and BIC scores for each model. Sum up the participant scores for each model (i.e. for each model you will have a single AIC score and a single BIC score). Report the results. Comment briefly. Which model would you choose as the "best" model?

For your reference, use the following equations for the calculations of AIC and BIC, where NLL is the negative log likelihood, p is the number of parameters, and n is the number of observations (240 in our experiment):

$$AIC = 2 \times NLL + 2 \times p \tag{7}$$

$$BIC = 2 \times NLL + p \times \log(n) \tag{8}$$

Task (j): Model recovery and confusion matrix (6)

We now want to check the reliability of our model comparison procedure using model recovery simulations. This will give us some indication about how much we can trust our results from Task (h). For each model, simulate data (using the functions you already implemented) multiple times. For each of these simulated data set, fit each model and use model comparison to choose the best model. Display your results in a confusion matrix. Comment briefly. (Make sure to explain your approach and results in appropriate detail.)

Task (k): Discussion (14)

Discuss the experiment and computational modelling analysis. Comment briefly on the order in which we performed this analysis. Did Tasks (g)–(j) change your interpretation of the results you made in Task (f)? What are your overall conclusions?

It might not be very realistic to ask participants (especially severely ill patients) to perform this task (of 240 trials) without breaks.

- Would it be okay to only collect half the trials?
- Alternatively, we could give them a short break every T trials (e.g. every 60 trials). Describe potential issues of this solution.
- Assume that for our experiment participants have to respond within a brief time window (otherwise the trial is skipped) and you know from experience that subjects will miss about 5% of responses due to time-outs. How and where would you incorporate this knowledge?

Assume our experiment also included punishments. For example, instead of a blank screen in case of a no-reward outcome a punishment was received. Discuss how the models (Task (h)) or their interpretation would change. Briefly describe possible additional hypothesis and how additional parameters might be able to address them. Explain how the relationship between Models 1 and 2 might change.

BONUS Task (l): Parameter constraints (3)

Re-implement and fit model 1 with (soft-)constrained parameters ($0 < \varepsilon < 1$ and $1 < \beta < 15$). Discuss your observations.

References

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., and Kusumi, I. (2015). Reinforcement learning in depression: a review of computational research. *Neuroscience & Biobehavioral Reviews*, 55:247–267.

Dombrovski, A. Y., Clark, L., Siegle, G. J., Butters, M. A., Ichikawa, N., Sahakian, B. J., and Szanto, K. (2010). Reward/punishment reversal learning in older suicide attempters. *American Journal of Psychiatry*, 167(6):699–707.

Edition, F. et al. (2013). Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc*, 21.

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, 134(6):1751–1764.

Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., and Steele, J. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain*, 131(8):2084–2093.

Ratcliff, R. and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological science*, 9(5):347–356.

Rock, P., Roiser, J., Riedel, W., and Blackwell, A. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological medicine*, 44(10):2029.

Smith, P. L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in neurosciences*, 27(3):161–168.