Practical: The EZ diffusion model

Leendert van Maanen*

translated to MATLAB by Ysbrand Galama

May 12, 2016

The code blocks below can be copy pasted onto the MATLAB command line. First, clear your workspace and make sure all the necessary functions available. Write your answers in a pdf file, include the necessary plots. You may write it as a report (good for your writing skills), however, answering each question within a section/enumeration is also accepted.

Take note: you do not need to write a lot of code, but running the full data can take some time, so test beforehand (on a small part of the data) if everything works!

Assignment 1

What are the differences between the EZ diffusion model and the "full" Ratcliff diffusion model?

Assignment 2

Assume you have a data set with the following properties

```
Pc = .9; % the proportion correct responses
MRT = 100; % the mean response time
VRT = 1000; % the variance in the response time
    distribution
```

the EZ diffusion parameters are provided by

```
>> [v,a,Ter] = EZdiffusionfit(Pc, VRT, MRT)

v =

    0.0145
```

*Reference: Wagenmakers, E.-J., van der Maas, H. L. J. & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. Psychonomic Bulletin & Review, 14, 3-22

```
a =

   1.5133


Ter =

  58.3080
```

with v: drift rate; a: boundary separation (threshold); Ter (non-decision time).

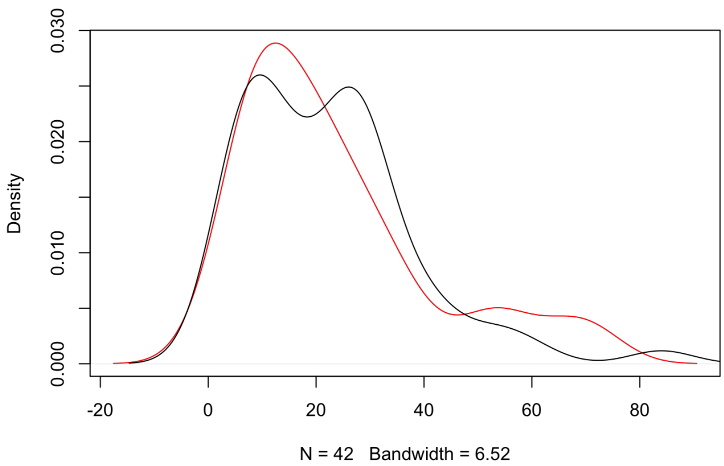
1. How do the parameters change if the proportion of correct responses changes?
2. How do the parameters change if the mean response time changes?
3. How do the parameters change if the variance in the response times changes?

Make a graph to illustrate the effects of behavioural changes on parameters.

Assignment 3

you have the following hypothetical data set:

```
dat = readdata('data1.txt');
correct = dat(:,1);
rt = dat(:,2);
[y1,x1] = ksdensity(rt(correct==0));
[y2,x2] = ksdensity(rt(correct==1));
plot(x1,y1,'-k',x2,y2,'-r');
```



Can you compute the EZ diffusion parameters (follow the steps below)? 1. inspect the data for suitability 2. Compute the relevant properties of the data 3. Compute the EZ diffusion parameters as above

Assignment 4

you have the following hypothetical data set:

```
dat = readdata('data2.txt');
correct = dat(:,1);
rt = dat(:,2);
```

1. Compute the EZ diffusion parameters.
2. Do these parameters make sense? If no, why not?

Assignment 5

you have the following data set, taken from a lexical decision experiment by Keuleers et al (2010):

```
loadBrysbaert
```

The data struct is called d, and the field names are (I think) self explanatory. Loading the dataset can take some time, so be patient.

1. Compute the EZ diffusion parameters per participant and condition (i.e., word or non-word, the wnw field)
2. Can you interpret the parameters? In other words, which (if any) of the parameters explains the difference between words and non-words?

3. Do these results make theoretical sense? Why (not)?

Assignment 6

The following two functions are needed. The first function (*ddiff.m*) concerns the joint density function of the diffusion model. The second function (*simdiff.m*) concerns a function that simulates data (rt's and accuracies) according to the diffusion model.

1. Simulate data for $N = 500$ trials with boundary separation, a, equal to 2, drift rate, v, equal to 1, and non-decision time, Ter, equal to 0.5 and plot the distributions. You can use the following code for generating data:

   ```
   rng(1310);
   a=2;
   v=1;
   N=500;
   ter=.75;
   [x,rt] =simdiff(N,a,v,ter);
   ```

2. Using the function ddiff above, determine the log-likelihood of the data for $a = 1$, $v = 2$, and $ter = 0.5$. You should obtain -2607.498

3. Use `MATLAB`'s build-in function *fminsearch* to maximize the log-likelihood of the unknown parameters a, v, and Ter. See *help fminsearch* for more details (BEWARE: *fminsearch* minimises the model, you want a maximum! (think how you can do this)). In *fminsearch* it is possible to provide upper and lower bounds, use $< 0, Inf >$ for a, use $< -Inf, Inf >$ for v, and $< 0, min(rt) - .01 >$ for Ter. Make sure that you use the data generated above (i.e., using rng(1310) (=seed)). The value of the likelihood function at its maximum should be -480.5922. As starting values $v = 1$, $a = 1$, and $Ter = 0.25$ worked for me.

4. Estimate the EZ-diffusion parameters using the function *EZdiffusion()*. Make sure that you use $s = 1$ in this function (the default is $s = .1$ which puts the parameters on a different scale). Compare these EZ-parameter estimates to the ones you obtained using maximum likelihood. What do you think of the similarity/differences?

Assignment 7

Simulate data using the following code

```
rng(1310);
a=2;
v=1;
N=500;
ter=.75;
[x1,tr1] = simdiff(N,a,v,ter);
a=1
[x2,rt2] = simdiff(N,a,v,ter);
rt= [rt1 rt2];
x= [x1 x2];
```

i.e., this code simulates data for 500 trials administered under an accuracy instruction ('answer as accurate as possible') and 500 trials administered under a speed instruction ('answer as fast as possible').

1. Which parameter is simulated to be affected by the manipulation?

We now introduce a new parameter Δa. This parameter models the difference between the a in the second condition as compared to the first condition. Thus it holds that

Accuracy condition (first 500 trials of vector rt)

- boundary separation = a
- drift rate = v
- non-decision time = Ter

Speed condition (first 500 trials of vector rt)

- boundary separation = $a + a$
- drift rate = v
- non-decision time = Ter

i.e., we have four parameters: a, Δa, v, Ter.

2. In the data as simulated above (using rng(1310)), what is the likelihood of the data for $a = 2$, $\Delta a = .5$, $v = 1$, and $Ter = .5$? You should obtain -1125.825.

3. Optimise the likelihood function for the unknown parameters (a, Δa, v, Ter). Use the same settings as before. For Δa, you can use $< -Inf, Inf >$ as boundaries. Make sure that you use the data generated above (i.e., using rng(1310)). The value of the likelihood function at its maximum should be -441.9329. As starting values $a = 1$, $\Delta a = .1$, $v = 1$, and $Ter = 0.5$ worked for me.

4. What do you think of the recovery of the true parameter values (i.e., are the parameter values that are used to create the data satisfactorily recovery in the parameter estimates)?

5. Do the above, but now for v. That is, simulate 500 trials with $v = 1$, $a = 1$, and $Ter = .75$ and 500 trials with $v = 2$, $a = 1$, and $Ter = .75$. Fit a model with an effect on v (i.e., introduce Δv similarly as above in case of a). If you use rng(1310) again in data generation, the maximum of your likelihood function should be 242.5705.

Assignment 8

In class we discussed signal detection theory, which transforms Hit rate and False Alarm rate in d' and c. Analogously, The EZ model transforms MRT, VRT and Pc in v, a, and Ter.

1. What are the conceptual similarities between SDT and EZ?

2. Does the criterion parameter c in SDT translate into the boundary parameter a in EZ? Why (not)?

3. Which parameter in the full DDM relates to the SDT criterion parameter?