Sequential Sampling Models for DFE

.

.

Draft of July 30, 2014

Basic model

The basic model is a sequential sampling model with two absorbing boundaries as shown in Figure **??**. Each trajectory (the colored curves) represents the process in time for choosing either option A or option B. It is continuous in time and does NOT represent the number of draws. The choice options are defined below.

For all model versions we assume at least two processes, one for choosing option A and one for choosing option B. The structure of the gambles is mapped onto the drift rate parameters μ_A and μ_B.

Structure of the gambles

I refer to Doug's output file.

Each of the gambles have a different expected value and a different variance.

- For the High (H) condition the expected value ranges from about 5.8 to 6.2 in the gain domain and from -4.8 to -5.2 in the loss domain.
- For the Low (L) condition the expected value ranges from about 4.8 to 5.2 in the gain domain and from -5.8 to -6.2 in the loss domain.
- The low variances range from about 10 to 30 and the high variance between 60 and 90.

- The difference in expected values between H and L ranges from .71 to 1.3.
- The values actually drawn vary between -24 and 24.

Except for the variances and the actual values the ranges are relatively small.

The question is what to map onto the drift rates. We want to weight the input (the parameters in the model) somehow. There are too many different expected values (altogether 166) to even consider this as an option. To consider the difference in expected values is not reasonable since both gambles are not directly compared (at least not for the low switchers but we will use that to construct the model for decisions from descriptions). And they would be too many as well (83). To use the actual values drawn could be one way to go. The ids for the set of outcomes are between 0 and 45 for the L condition and between 2 and 45 between the H condition. That is we have max 46 different values (hopefully less).

Specific models

Using the actual values drawn gives one possible input for the L condition to model:

$$\mu_{L_{id_n}} = (w_i L_{x1n} + w_j L_{x2n} + w_k L_{x3n})/\sqrt{Var(L_n)} \tag{1}$$

The same holds for the H condition. Implicitly we assume a zero input for H when L is considered and zero input for L when H is considered. w_i, w_j, and w_k are the parameters to be estimated from the data. With max 46 different values this results in 46 different w.

An alternative input is

$$\mu_{L_{id_n}} = (w_i L_{x1n} + w_j L_{x2n} + w_k L_{x3n})/para(Var(L)) \tag{2}$$

The difference between both inputs is that in Eq. 1 the standard deviation from the actual gambles are taken whereas in 2 low and high variances build two clusters and each is presented by an additional parameter. There are certainly more ways to set up the drift rate, e.g., by combining the values in the gambles in a reasonable way or by finding a different model input structure. We should discuss it.

Whatever the input to the drift is there are also various model architectures we have discussed briefly (there are certainly more, but let's start with these). For simplicity the processes are shown in the following without the trajectories.

Model 1

We assume one process for both options H and L with two subprocesses. In one subprocess H is considered and in the other subprocess L is considered.

t_1 indicates when the DM switched from H, say, to L. In the model t_1 is a RV (e.g., Uniform with small variance, i.e., almost fixed or with high variance which approaches a geometric distribution) and t_1 may be even related to the average number of draws. Low switchers switching happens once; f high switchers often. This is mapped one to one to the process: for low switchers we assume one switch and for frequent switchers the processes go back and forth. The sequence in which the processes are considered (first H and the L or vice versa) matters for the prediction (not for the choice probabilities but for the mean RT). We could also assume that the DM start randomly with H or L. Note that the starting process of the second process is not exactly as suggested in the figure. The trajectories are random (see **??** and so are the next starting positions.

Model 2

The second model assumes two separate processes, one for H and one for L. However, they are not directly interlinked. Note that the choice options are now accept/reject H and accept/reject H. Also note, that the boundaries may be different in size.

Both processes are combined by assuming a mixture of processes. With probability p H is processes and with probability $(1 - p)$ L is processed. p is a model parameter but may be directly linked to the number of draws taken from each option. I don't know yet how to model the switching behavior for this model.

Remarks

- Most importantly to get it started is to decide what features of the gambles should be map onto the model parameters.
- Risk attitude can be mapped onto a approach/avoidance parameter like in DFT (cf. forgetting). That is, the drift rate in Eqs. 1 and 1 are extended by $-\gamma x$, where γ is a model parameter and x is a state in the state space (roughly the evidence to accumulate).

Exploratory analysis

Additional analyses were conducted to answer the following questions:

- Are overall measures like sample size and number of switches related to subject- and problem-level factors?
- Can trial-by-trial decisions (whether to continue sampling the same option, to switch to the other option, or to stop sampling altogether) be predicted by changes in the evidence accumulated?

Subject- and problem-level analysis

The following predictors were considered:

- Group (young/old)
- Session (1-21)
- Domain (gain/loss)
- Pairtype (HH-LL or HL-LH): all option pairs included one high-variance and one low-variance option. For HH-LL pairs, the H option had high variance (this was predicted to be a difficult problem); in HL-LH pairs, the H option had low variance (predicted to be an easy problem).
- EV-diff: difference in expected value between H and L option
- Total variance (combined between two options; although all pairs had a high and low variance option, there was still some variability in the total variance across different pairs)

Number of switches. Ran mixed effects linear regression on the number of switches, with the variables above as fixed effects, and a random effect term for sample size by participant. This random effect leads to a different coefficient for each participant describing the effect of samples on the number of switches. For frequent switchers this value is high (since number of switches is highly correlated with number of samples), while for people who switch the same number of times regardless of sample size it should be at zero.

There were no significant effects of any of the variables above on the number of switches. This result was consistent with earlier analyses that indicated 1) a bimodal distribution of switching rates, and 2) stable switching rates within subjects. For following analyses, people were divided into two groups based on their switching rate (FREQ and RARE switchers) using a median split on the average switching rate.

Sample size. The next question was how the total sample size was related to the same set of predictors (now adding the switching group variable based on median split). There was a significant effect of the switch group (FREQ/RARE; $p < .001$), with frequent switchers taking fewer samples than rare switchers. There was also an effect of session ($p < .001$), with the number of samples decreasing over time. Additional modeling indicated that this decrease in sample size was specific to the RARE group. However, there were no other effects of group or problem on sample size.

In sum, the results suggest that after accounting for variability between participants, the features of the problem did not have a significant impact on switching rate or overall sample size.

Trial-by-trial analysis

Although the preceding analysis didn't reveal any significant relationships between features of the problem and sample size or overall switching rate, it may be that trial-by-trial changes in a participant's experience were related to when they made decisions to switch or stop.

The following trial-by-trial variables were considered:

- Sample outcome
- Deviation of outcome from sample mean
- Absolute deviation from sample mean
- Deviation from streak mean
- Absolute deviation from streak mean
- Difference in sample means (MN-DIFF)
- Sample variance

Stay vs. leave. The first set of models compared decisions to STAY (continue sampling the same option) or LEAVE (*either* switch to a different option or stop sampling). Since frequent switchers chose to LEAVE on nearly every trial, here the focus was on rare switchers' decisions to stop sampling the current option.

There was a significant effect of absolute deviation ($z = 10.1$, $p < .001$), with LEAVE decisions more likely following larger deviations from the sample mean. In addition, there was a significant effect of sample variance ($z = -10.2$, $p < .001$), with LEAVE decisions more likely as sample variance decreased.

Switch vs. stop. The second set of models focused on the subset of trials in which LEAVE decisions were made, comparing SWITCH and STOP decisions. Separate models were fit for RARE and FREQ groups of participants.

For rare switchers, there was a significant effect of MN-DIFF ($z = 3.4$, $p < .001$), with STOP decisions more likely as the difference between sample means increased. There was also a significant effect of pooled variance ($z = -11$, $p < .001$) with STOP decisions more likely with decreased variance. Finally, there was a small effect of deviation from the sample mean ($z = 2.1$, $p = .03$).

For frequent switchers, STOP decisions were more strongly predicted by the outcome of the last sample. There were significant effects of the deviation of the last outcome from the sample mean, $z = 6.6$, $p < .001$ (a similar result is found when using absolute deviation). There was no effect of the difference in sample means, but there was a significant effect of pooled variance ($z = -19$, $p < .001$).