Slot-machine purchasing task:
Proposal for Pilot 2 for the dynamic allocation of memory resources project

Antonio Rangel
2022 07 23

PRELIMINARY COMMENTS

Before describing the proposed task for Pilot 2, let me start with some background about my thinking:

- I view the key goal of the experiment to test the key idea of the DRA theory; namely, than when subjects learn the values of action-state pairs in a RL context, their allocation of mental resources (and thus their precision of the learnt representations) depends on the expected return/value of the those representations.
- In particular, I view the theory as predicting that memories for action-value states that are encountered more frequently or in decisions involving higher stakes (i.e., where the consequences of mistakes are larger), should be encoded more precisely.

- The primary goal of a first experiment on this should be to carry out the simplest and cleanest possible test of this qualitative predictions.
- Testing the precise quantitative predictions of the model is desirable, but of secondary importance in the first experiment.

- There are some challenges in testing the theory:
    - The model utilizes very large numbers of repeated trials in relatively simple settings and assumes that the algorithms/processes used by the subject during the experiment remain unchanged throughout the task.
    - The later assumption is unlikely to hold in many naturalistic settings and experiments.
    - Consider the grid world in the examples in the paper. After enough trials, the animal can just switch to more "semantic style" representations like go "down-down-left" or "head north-west" instead of sampling and comparing values.
    - In fact, I now think that this was a key problem in the first pilot.

- The goal of the new experiment proposed below is multi-fold:
    - Distill the task to the simplest possible essentials.
    - Create a setting in which subjects maintain a constant value sampling and comparison process through the entire task to avoid the challenge above.
    - Design a task that can be carried out in a single day of data collection to minimize subject recruitment and data quality problems
    - Design a task that provides very simple and direct tests of the theory.

TASK DESCRIPTION

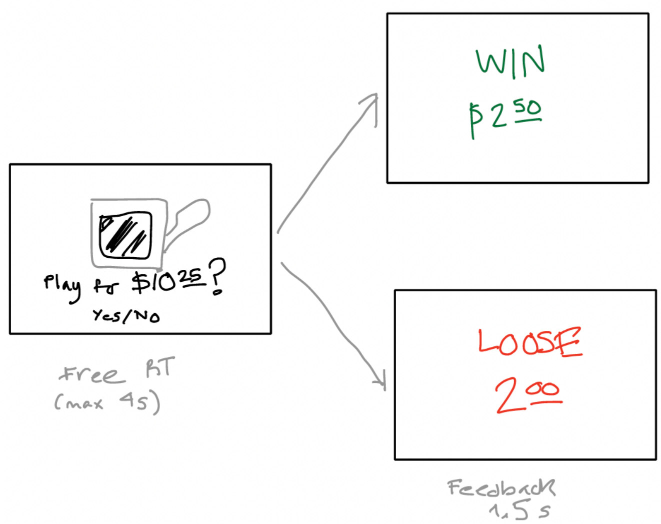
Every trial, the subject is shown one of 4 different slot-machines as well a price. Let s_t and p_t denote, respectively, the slot-machine and the price shown in trial t. The subject needs to decide if he wants to purchase the right to play the lottery by paying the price p_t: Yes (by pressing left arrow) or No (by pressing right arrow).

Each trial, the reward of playing the lottery is a Gaussian random variable with a fixed but unknown mean for each slot-machine, and a constant variance $\bar{\sigma}^2$ for all slot machines.

If he chooses No, the reward for the trial, denoted r_t, is 0.

If he chooses Yes, the reward for the trial is a Gaussian random variable with mean $\mu(s_t) - p_t$ and variance $\bar{\sigma}^2 = 25$, where $\mu(s_t)$ denotes the mean of the slot-machine shown in the trial. Rewards are rounded to the nearest cent.

The timing and look of each trial is as in this figure:

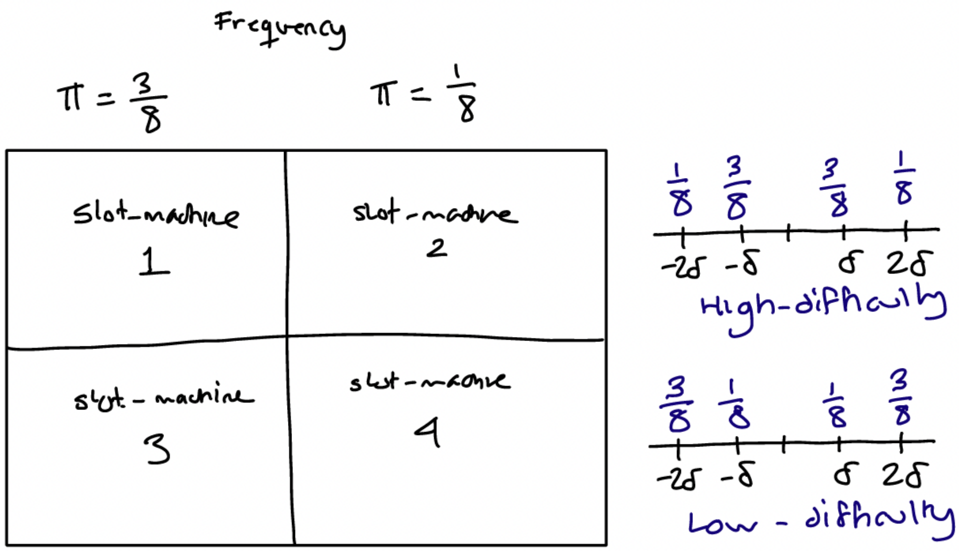


Note:

- If the subject chooses No, there is no Feedback screen
- Trials separated by a 500 ms blank screen
- If the subjects does not make a choice within 4 s, show a screen saying “please respond faster” for 1.5 s.

There mean of the slot-machines is randomly drawn uniformly from $5 to $15, in 50 cent increments. Subjects are told this during the instructions, but not the actual values of the slot-machines.

The slot-machines differ in two dimensions: (1) the (independent) probability with which they are shown in each trial, and (2) the difficulty of the decisions associated with them. The structure of the slot-machines and their associated trials is as follows:



I propose starting the pilot with a value of $\delta = \$0.75$

Let d_t denote the difficulty selected for trial t, which is sampled based on the distribution of difficulties described for each machine. Then the price shown in trial *t* is $p_t = \mu(m_t) - d_t$.

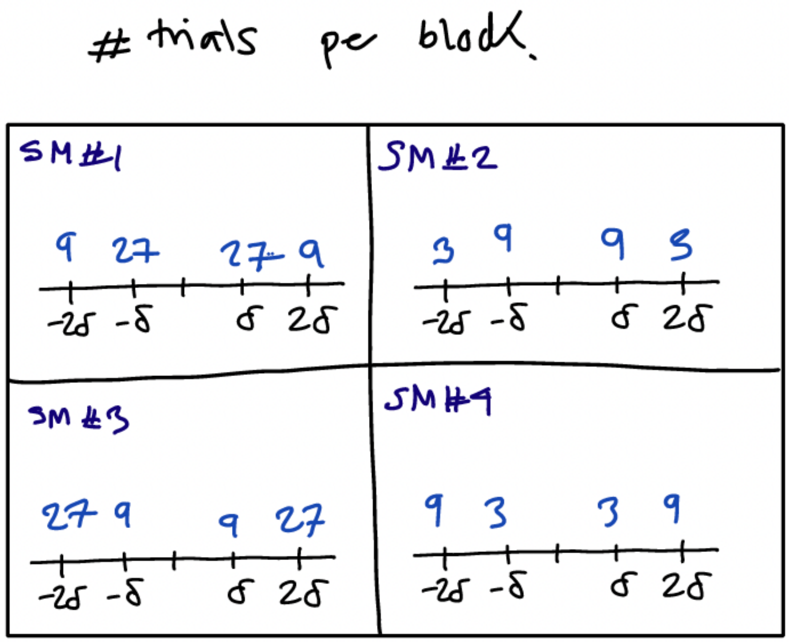
Note a few things about the task structure:

- There is a 3-to-1 ratio between the high and low frequency fractals.
- There is also a 3-to-1 ration between the probability of high difficulty (with $d_t = \delta, -\delta$) or low difficulty (with $d_t = 2\delta, -2\delta$).

The experiment is structured as follows:

- Subjects are given written task instructions.
- Subjects participate in 4 blocks of 184 trials (with an estimated duration of 9-16 mins p/ block)
- Blocks 1 and 2 will be used as a learning bock.
- Blocks 3 and 4 will be used as the post-learning test data.

- Each block is semi-randomized so that the trial count match the figure below. There is also a constraint that the same slot-machine not be shown in two consecutive trials.
- Subject compensation for the task is determined and has two components:
  - A guaranteed show-up fee of approx. $30 p/ hour.
  - The sum of earnings in two random trials from each block, capped below at zero.

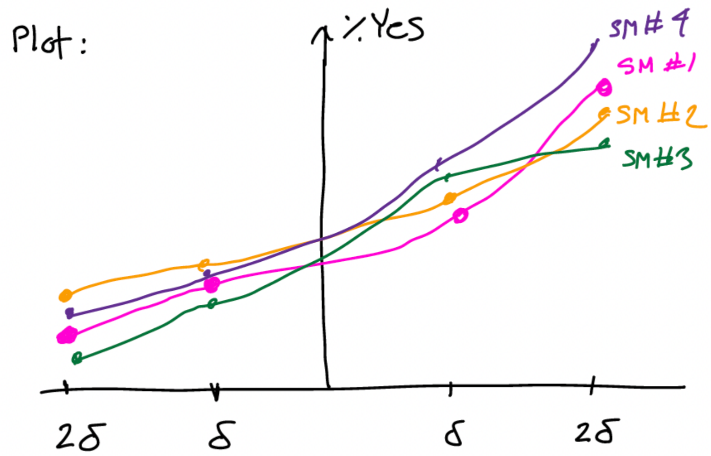


PROPOSED TESTS

The key data for the experiment are the number of Yes/No choices for each fractal and level of difficulty $(-2\delta, -\delta, \delta, 2\delta)$, in the two post-learning test blocks.

This choice data can be analyzed in two natural ways.

Basic test. As the following shows, we can plot the choice curves for each of the four fractals. (This should be done in two steps: (1) compute the mean %Yes for each subject and bucket, then compute the group mean of means, and add an error bar with width given by the standard deviation of the individual means).



This provides a basic test of the model:

- In the equal precision case, the four curves should lie on top of each other.
- In the case in which precision depends only on frequency, the curves for SMs 1 and 3 should be identical and steeper than the curves for SMs 2 and 4 (which are also indistinguishable from each other)
- In the case in which precision only depends on difficulty, the analogue should hold for SMs 1 and 2 vs SMs 2 and 4.
- In the DRA case, the four choice curves should separate, and their relative steepness is as predicted by the Patel-Acerbi-Pouget DRA model.

Why is this the case?

- Suppose, as assumed in the theory, that decisions are made by drawing a sample from a Gaussian with the true mean (which has been learnt by RL) and a precision determined by the model.
- Then the probability of choosing Yes in any trial is given by a standard Gaussian: $P(Yes) = \Phi(\frac{d_t}{\sigma(m_t)})$, where $\sigma(m_t)$ is the precision assigned to slot-machine m_t.
- It follows that in each bucket d_t the slope of the choice curve is determined solely by differences in the precision.
- The same holds if choices follow a Logit instead of Probit model.

Better test. A better quantitative test of the model involves estimating a hierarchical logit of probit model of the choice data.

RT test. It would also be good to do the RT version of the choice figure above to check that there are no differences across slot-machines, so that the results are driven by differences in encoding.

NOTES: POTENTIAL TASK IMPROVEMENTS

- Fine tune the key parameters δ, $\bar{\sigma}^2$ and the distribution of difficulties/
- Use 8 fractals
- Use graphical representation for prices and payoffs, instead of numbers.