Uncertainty and Exploration in a Restless Bandit Problem

Midway Progress Report

Harsh Arora (20218266)   Pratham Shukla (17816497)

1   Progress

The paper by Speekenbrink and Konstantinidis (2015) investigates uncertainty and exploration in a four-armed restless bandit task using a set of four conditions for the bandit arms. The conditions are constructed using a 2X2 setting of stable-variable and trend-no trend features. The authors use a number of models using combinations of learning and choice rules to fit the data. We use only the best performing models for each learning rule for the project, which are,

1. Kalman Filter and probability of maximum utility

2. Model-free decay rule and fixed temperature softmax

3. Model-free delta rule and dynamic temperature softmax

The data will be fitted using log-likelihood maximization with the Nelder-Mead Simplex algorithm. An illustration of the Kalman Filter - Fixed Softmax model has been added in this notebook.
The files pertaining to the experiment can be found here. The experiment will be live for the class soon.

2   Expected Results

As mentioned in the paper, we expect to find $\triangle(AIC)$ and $\triangle(BIC)$ scores for the three models above. The model with the Kalman Filter for a learning rule is expected to fit the most number of participants. The models shall also be tested for degrees of risk-aversion and loss-aversion using the definition of utility from the Prospect Theory (Tversky & Kahneman, 1992, Ahn et al., 2008).

To study the behavioral results, a generalized mixed-effects model shall be used. Following the paper, the 200 trials will be divided into four blocks of 50. The performance is expected to be explained by significant volatility-block and trend-block interactions, since both, volatility and trend closely determine the discriminability of the blocks. It'll be interesting to revisit these results since the authors used a different set of rewards instantiations.

A similar analysis of the switching behavior shall also depict volatility-block, trend-block interactions, and three-way trend-volatility-block interactions.

The specifics of each of these interactions will be covered in the final report.

3   References

1. Speekenbrink M, Konstantinidis E. Uncertainty and exploration in a restless bandit problem. Top Cogn Sci. 2015 Apr;7(2):351-67. doi: 10.1111/tops.12145. Epub 2015 Apr 20. PMID: 25899069.

2. Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty, 5, 297–323.

3. Ahn WY, Busemeyer JR, Wagenmakers EJ, Stout JC. Comparison of decision learning models using the generalization criterion method. Cogn Sci. 2008 Dec;32(8):1376-402. doi: 10.1080/03640210802352929. PMID: 21585458.