Revisiting the multi-armed bandit problem

Jeremie Houssineau, Department of Statistics

jeremie.houssineau@warwick.ac.uk

Background: It is commonly accepted in statistics that the only way of dealing with uncertainty is through the mathematical tools characterising random phenomena. However, it is often information, or rather the lack of it, that is the source of uncertainty. Using probability theory in this case has known limitations such as the absence of non-informative priors in unbounded spaces. Recent research [1, 2] however suggests that the measure-theoretic concept of outer measure can be used to model the lack of information about some parameters in a statistical model and to derive practical algorithms that complement the existing wealth of statistical techniques in a principled and yet intuitive way.

Description: The multi-armed bandit problem describes the situation in which an agent has to decide at each time step between exploration (trying a new slot machine) and exploitation (using the slot machine that is believed to have the best chance of reward according to the available information). The most widespread solution to this problem is referred to as Thompson sampling [3] and is based on the selection of the action that maximises the expected reward. Although intuitive, Thompson sampling remains a heuristic.

The type of outer measures that has been considered in [1] relies on a mixture between maxima and integrals and allows for representing a complete lack of information a priori. There is therefore a strong connection with the exploration-exploitation dilemma where the probability of different rewards is typically unknown a priori and where both expectation and maximisation appear. The objective of the project is to formalise this connection and use the mechanisms associated with outer measures to explore other solutions to this problem.

References

[1] Jeremie Houssineau, Neil K Chada, and Emmanuel Delande. Elements of asymptotic theory with outer probability measures. *arXiv preprint arXiv:1908.04331*, 2019.

[2] Jeremie Houssineau. Parameter estimation with a class of outer probability measures. *arXiv preprint arXiv:1801.00569*, 2018.

[3] William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.