Statistical properties of decision environments

Özgür Şimşek
Max Planck Institute for Human Development
ozgur@mpib-berlin.mpg.de

There is evidence that people often make decisions by using simple rules, such as lexicographic heuristics or unit weighting. These decisions can potentially have low accuracy because they use few pieces of information and do not try to combine the information optimally. Surprisingly, empirical studies have shown the opposite: simple rules often perform almost as well as or better than statistical decision methods, including regression models, decision trees, and support vector machines.

Several properties of the decision environment can explain this finding, including for example, dominance relationships among decision alternatives and noncompensatoriness in the environment (more specifically, in the relationship between the decision cues and the decision criterion). When these properties are present, the decision problem becomes easy in the sense that many methods (simple and complex) yield the same decision, eliminating the need to find an optimal trade-off between the various cues that are available. For instance, using only a single cue, or a simple tallying of the cues, may suffice to make the correct decision.

In this talk, I will review statistical properties of the decision environments that support simple decision rules by making the decision problem easy. I will then examine how often we encounter these properties in a large collection of natural environments, obtained from a wide variety of sources, including data repositories on the web, textbooks, research publications, packages for R statistical software, and individual scientists collecting field data, on diverse subjects, including biology, business, computer science, ecology, economics, education, engineering, environmental science, health, medicine, political science, psychology, sociology, sports, and transportation. The main results is that these properties are highly prevalent, making it possible for simple rules to reach (and even exceed) the predictive accuracy of statistical models, using less information and less computation.

The properties I will review all affect the bias component of the prediction error. The results imply that, while decision heuristics examine a tiny fraction of the space of possible decision models, they may do so without introducing much additional bias. This takes the bias/variance discussion in this field of research, which has so

far focused on how decision heuristics reduce variance at the expense of increasing bias, to a different direction: decision heuristics may incur low bias in addition to low variance.

I hope that these results will stimulate further research in statistical properties of natural decision environments and in developing models of bounded rationality that exploit them.

This talk will complement the talk by Konstantinos Katsikopoulos, who will present a broad conceptual overview of rationality of simple decision models.