Agent Based Models and Sensitivity Analysis – Professor Allie Lewis

Outline: Agent-based models (ABMs) are capable of simulating complex biological phenomena at the microscale level by tracking individual agents whose behavior may be governed by a variety of different mechanisms. However, this level of model complexity comes at a cost; ABMs are often computationally expensive, such that a single simulation may take hours or even days to run. Due to this cost, many researchers forgo conducting adequate sensitivity analyses on ABMs, even when such analyses are vital to understanding the robustness of the model.

Borgonovo et. al (2022) recently outlined a procedure for conducting a rigorous sensitivity analysis on ABMs. However, the proposed procedure may still be highly limited by computational budget unless we can create a more intelligent sampling plan, in the sense that we limit model evaluations to input regimes in which the quantity of interest is likely to be highly sensitive to fluctuations in the parameter values. One commonly employed method for global sensitivity analysis is Morris screening (Morris, 1991), in which inputs are ranked according to their sensitivity by averaging over coarse, local derivatives. As detailed in Lewis et. al (2016), the use of active subspace methodology to exploit linear combinations of parameters which are most influential to the quantity of interest may be combined with the Morris sensitivity metric in order to improve the sampling mechanism. By applying such a sampling strategy to the sensitivity analysis procedure for an agent-based model, we can make intelligent choices about how to spend our limited computational budget.

Research objectives: To enable more efficient sensitivity analyses of ABMs by aligning the sampling strategy according to the most influential directions in the input space. Computational cost and feasibility will be compared with previously employed methods to illustrate that the proposed sampling scheme can produce the same sensitivity results with fewer model evaluations.

Outcomes: Active subspaces are one of an emerging set of tools for dimension reduction in the field of uncertainty quantification (Constantine, 2015). Students on this project will explore the tradeoffs between model complexity and computational costs, and will learn how to apply techniques of dimension reduction to allow for robust model analyses. In addition to developing and enhancing skills in linear algebra and probability, students will gain extensive experience in writing simulation code.

Pre-requisites: Some experience with linear algebra and/or calculus-based probability recommended. Programming experience preferred.

References:

[1] E. Borgonovo, M. Pangallo, J. Rivkin, L. Rizzo, and N. Siggelkow, “Sensitivity analysis of agent-based models: a new protocol,” *Computational and Mathematical Organization Theory*, Vol. 28, pp. 52-94, 2022.

[2] M.D. Morris, “Factorial sampling plans for preliminary computational experiments,” *Technometrics*, Vol. 33(2), pp. 161-174, 1991.

[3] A.L. Lewis, R.C. Smith, and B.J. Williams, “Gradient-free active subspace construction using Morris screening elementary effects,” *Computers and Mathematics with Applications*, Vol. 72(6), pp. 1603-1615, 2016.

[4] P.G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies,* SIAM Spotlights, Philadelphia, PA, 2015.

Orderings of Point Sets – Professor Carl Hammarsten

Outline: Consider a set *S* of *n* points in the plane. A choice of a vantage point *p* in the plane then defines an ordering *O* on *S* according to each point's distance from *p*. For $n > 3$, only a specific subset of the $n!$ possible orderings are achievable in this way. Good and Tideman have shown the maximum number of attainable orderings is a sum of unsigned Sterling numbers of the first kind. Carbonero, et al, have shown the minimum number of attainable orderings is $2n - 2$. They have also constructed a complete list of possibilities for the $n = 4$ case and derived a partial list of restrictions for the lists of possibilities when $n > 4$. In yet-to-be-published work from last summer, another REU group derived the full list of possibilities for $n = 5$ and developed techniques that should ideally extend to higher values of n.

Research Objectives: We will endeavor to find complete descriptions of the achievable number of orders for $n > 5$. As part of this goal, we will examine the related, though somewhat more general, question arising from considering a set *S*, a vantage point *p*, and a desired ordering *O*, in which case we ask what is the minimal modification to a point in (or subset of) *S* required to make *O* achievable.

Outcomes: Everyone in this group will play a major role in all of our work and hopefully feel equal ownership from the beginning. This project has a nice balance of identifying theoretical requirements and balancing those with constructive examples. Approaching the list of the possible number of orderings from both angles requires both strong proof writing technique and the ability to create unexpected examples.

The nature of this project also leads to many chances to modify and generalize ideas to different settings – i.e. by changing the ambient space or the number of observers. Students will thus have ample opportunity to experience the ever-changing experience of "research".

Prerequisites: Some experience with linear algebra and/or discrete mathematics recommended.

References:

Carbonero, A., Castellano, B., Gordon, G., Kulick, C., Schmitz, K., & Shelton, B. (2021). Permutations of point sets in R^d.

Good, I. & Tideman, T. (1977). Stirling Numbers and a Geometric Structure from Voting Theory. J. Comb. Theory, Ser. A. 23. 34-45. 10.1016/0097-3165(77)90077-2.

Bayesian Applications to Polygraph Testing – Professor Jeff Liebner

The Empirical Scoring System (ESS) is a mathematical technique that is used by field examiners working with polygraphs as a scoring model based on simple measurements. The promoters of the technique argue that the simplicity of the scoring method means that those who are both experienced or inexperienced with mathematical techniques and polygraphs can obtain results that are easily interpretable.

The idea behind polygraphs is that bodily responses can be measured when a subject is exposed to an examination and responds with a falsehood or a true statement. Common measurements include examining the subject's respiration amplitude, respiration rate, changes in respiratory baseline, presence of apnea, electrodermal activity, and cardiovascular activity, among others. According to the ESS model, each measurement is scored as a +1, 0, or -1, and a composite score is used to assess whether or not the subject is lying.

The Empirical Scoring System-Multinomial (ESS-M) was an update to the ESS model that employed what the creators described as a Bayesian model. The argument is that the posterior odds obtained from the model are more intuitive than the p-values obtained from the traditional frequentist p-values. This model uses a prior that assumes that a subject is equally likely to be telling the truth as telling a lie and uses a likelihood that is built around equally likely probabilities for the three outcomes of +1, 0, and -1 for the measurements of the subject's bodily responses. Despite its widespread use in the polygraph community, this update has serious mathematical flaws, especially given its claims to be a Bayesian model.

Research Objectives: To examine the mathematical structure behind the ESS-M model and identify its shortcomings. After identifying these problems, the team will develop its own Bayesian model or models for assessing polygraph results.

Outcomes: Bayesian techniques are a growing field in statistical research and machine learning. Participants will develop a conception of these tools to broaden their mathematical and statistical expertise. Contacts within the polygraph community will lead to a spread of the ideas developed in this research project, correcting existing false assumptions and promoting a more useful and accurate tool for this community. Thus, participants will be coached to communicate mathematical results to those with less mathematical training.

References:

Nelson, R. (2017). Multinomial reference distributions for the Empirical Scoring System. Polygraph & Forensic Credibility Assessment, 2017, 81-115.

Nelson, R., Handler, M., Shaw, P., Gougler, M., Blalock, B., Russell, C., Cushman, B. & Oelrich, M. (2011). Using the Empirical Scoring System. Polygraph, 40, 67-78.

Nelson, R. (n.d.) *Lafayette Tech Talk: Understanding the Lafayette ESS-M Narrative Summary*. Lafayette Polygraph. Retrieved February 22, 2023 from https://lafayettepolygraph.com/downloads/AAPP-Lafayette-ESS-M-Narrative-Summary.pdf

Nelson, R. (2018). Five Minute Science Lesson: Bayes' Theorem and Bayesian Analysis. APA Magazine, 51, 65-78.