INVENTEURS DU MONDE NUMÉRIQUE

PhD topic:
(Optimal) adaptive testing for clinical trials

Context This PhD is founded by Inria through an "Action Exploration" (AEx) program (BETA-3K). The PhD candidate will work in the Scool team of Inria Lille under the supervision of Emilie Kaufmann (CNRS, https://emiliekaufmann.github.io/) and Rémy Degenne (Inria, https://remydegenne.github.io/). Part of the research will also be done in collaboration with the team of Sofia Villar (MCR Biostatistics Unit, University of Cambridge).

Research topic A multi-armed bandit is a simple model for sequential resource allocation tasks: an agent is facing different options, called arms, with random outcomes (sometimes called rewards), and the goal is to sequentially select these arms in order to achieve a certain objective, such as maximizing the (expected) sum of the outcomes or identifying the arm with largest mean outcome [Lattimore and Szepesvari, 2019]. The introduction of bandit models dates back to the 1950s [Robbins, 1952] as a simple model for an *adaptive* clinical trial: out of K treatments on trial, a doctor can select for each patient which treatment to try, based on the results from the past trials. This flexibility has a potential to reduce the number of patients allocated to less efficient treatments or to help identify the best treatment faster. Despite this, multi-armed bandit algorithms, mostly developed in the machine learning community have been seldom used in clinical trials, for different reasons [Réda et al., 2020, Robertson et al., 2023].

One of this reason is that for confirmatory trials, in which the new candidate treatment is typically compared to the standard-of-care or to a placebo, there is a well established statistical methodology. The requirement is often to build a statistical test to reject the hypothesis ($p_{\text{new}} = p_{\text{old}}$) with a guaranteed type I error, and the number of patients assigned to each treatment is chosen to guarantee some power against some alternative $p_{\text{new}} > p_{\text{old}} + \varepsilon$. Classical (batch) statistical tests are used and their calibration require the treatment allocation to be fixed in advance. It has been observed that using these tests with data that has instead been collected adaptively can inflate the type I error. For example, the work [Williams et al., 2021] presents a simulation study showing what can happen when combining a Wald test with data collected using Thompson Sampling, a particular bandit algorithm. Still, the use of other type of tests may be better suited for a combination with adaptively collected data. One of the goals of this internship is to perform a thorough numerical comparison of different combinations of bandit algorithms and tests, on realistic benchmarks, and compare them with different strategies used in the literature on adaptive clinical trials. We will for example investigate the use of Generalized Likelihood Ratio Tests, that can be calibrated even with adaptively collected data

CENTRE DE RECHERCHE
LILLE - NORD EUROPE
Parc scientifique de la Haute-Borne
40 avenue Halley - Bât A - Park Plaza
59650 Villeneuve d'Ascq - France
Tél. : +33 (0)3 59 57 78 00
Fax : +33 (0)3 59 57 78 51
www.inria.fr

[Kaufmann and Koolen, 2021]. As for the bandit algorithms, we will go beyond Thompson Sampling which is designed for maximizing rewards, and investigate recent algorithms from the best arm identification literature [Jourdan et al., 2022, Jourdan et al., 2023].

Interestingly, both the adaptive clinical trial and the bandit literature have considered some notions of an "optimal allocation" over the arms (with different definitions, see, e.g. [Rosenberger and Ricks, 2001, Garivier and Kaufmann, 2016]) and have derived adaptive sampling strategy for which the empirical allocation matches those, using a so-called Targetting [Hu and Zhang, 2004] or Tracking [Garivier and Kaufmann, 2016] approach. While in the bandit literature, the optimal allocation for fixed-confidence identification is well defined for any number of arms K [Garivier and Kaufmann, 2016], optimal allocation in the clinical trials literature are mostly defined for two-armed scenarios (that are the most common in confirmatory trials). The reason is that defining a testing objective in less clear in the presence of more than 2 hypotheses. Still, some attempt of defining optimal proportions in the K-armed scenario have been made [Tymofyeyev et al., 2007, Biswas et al., 2011]. We will review existing approaches, possibly find new ones and compare them to the bandit optimal proportions. The goal of this study is to find an adaptive allocation strategy that performs well, theoretically and empirically, both from the bandit "best arm identification" perspective and from an appropriate testing perspective in a clinical trial with potentially more than 2 arms.

Requirements The candidate should have a good background in statistics and machine learning. Some knowledge on multi-armed bandit algorithms is a plus, as well as coding skills in Python and Julia.

References

[Biswas et al., 2011] Biswas, A., Mandal, S., and Bhattacharya, R. (2011). Multi-treatment optimal response-adaptive designs for phase iii clinical trials. *Journal of the Korean Statistical Society*, 40(1):33–44.

[Garivier and Kaufmann, 2016] Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference On Learning Theory*.

[Hu and Zhang, 2004] Hu, F. and Zhang, L.-X. (2004). Asymptotic proerties of doubly adaptive biased coin designs for multitreatment clinical trials. *Annals of Statistics*, 32(1).

[Jourdan et al., 2022] Jourdan, M., Degenne, R., Baudry, D., de Heide, R., and Kaufmann, E. (2022). Top two algorithms revisited. In *Advances in Neural Information Processing Systems (NeurIPS)*.

[Jourdan et al., 2023] Jourdan, M., Degenne, R., and Kaufmann, E. (2023). An ε-best-arm identification algorithm for fixed confidence and beyond. In *Advances in Neural Information Processing Systems (NeurIPS)*.

CENTRE DE RECHERCHE
LILLE - NORD EUROPE
Parc scientifique de la Haute-Borne
40 avenue Halley - Bât A - Park Plaza
59650 Villeneuve d'Ascq - France
Tél. : +33 (0)3 59 57 78 00
Fax : +33 (0)3 59 57 78 51
www.inria.fr

[Kaufmann and Koolen, 2021] Kaufmann, E. and Koolen, W. (2021). Mixture martingales revisited with applications to sequential tests and confidence intervals. *Journal of Machine Learning Research*, 22(246).

[Lattimore and Szepesvari, 2019] Lattimore, T. and Szepesvari, C. (2019). *Bandit Algorithms*. Cambridge University Press.

[Robbins, 1952] Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535.

[Robertson et al., 2023] Robertson, D. S., Lee, K. M., López-Kolkovska, B. C., and Villar, S. S. (2023). Response-adaptive randomization in clinical trials: from myths to practical considerations. *Statistical Science*, 38(2).

[Rosenberger and Ricks, 2001] Rosenberger, W.F., S.-N. I. A. H. C. and Ricks, M. (2001). Optimal adaptive designs for binary response trials. *Biometrics*, 57:909–913.

[Réda et al., 2020] Réda, C., Kaufmann, E., and Delahaye-Duriez, A. (2020). Machine learning applications in drug development. *Computational and Structural Biotechnology Journal*, 18:241–252.

[Tymofyeyev et al., 2007] Tymofyeyev, Y., Rosenberger, W. F., and Hu, F. (2007). Implementing optimal allocation in sequential binary response experiments. *Journal of the American Statistical Association*, 102:224–234.

[Williams et al., 2021] Williams, J. J., Nogas, J., Deliu, N., Shaikh, H., Villar, S. S., Durand, A., and Rafferty, A. (2021). Challenges in statistical analysis of data collected by a bandit algorithm: An empirical exploration in applications to adaptively randomized experiments. *arXiv:2103.12198*.

CENTRE DE RECHERCHE
LILLE - NORD EUROPE
Parc scientifique de la Haute-Borne
40 avenue Halley - Bât A - Park Plaza
59650 Villeneuve d'Ascq - France
Tél. : +33 (0)3 59 57 78 00
Fax : +33 (0)3 59 57 78 51

www.inria.fr