Online Local False Discovery Rate Control: A Resource Allocation Approach

We consider the problem of sequentially conducting multiple experiments where each experiment corresponds to a hypothesis testing task. At each time point, the experimenter must make an irrevocable decision of whether to reject the null hypothesis (or equivalently claim a discovery) before the next experimental result arrives. The goal is to maximize the number of discoveries while maintaining a low error rate at all time points measured by local False Discovery Rate (FDR). We formulate the problem as an online knapsack problem with exogenous random budget replenishment. We start with general arrival distributions and show that a simple policy achieves a $O(\sqrt{T})$ regret. We complement the result by showing that such regret rate is in general not improvable. We then shift our focus to discrete arrival distributions. We find that many existing re-solving heuristics in the online resource allocation literature, albeit achieve bounded loss in canonical settings, may incur a $\Omega(\sqrt{T})$ or even a ￼ regret. With the observation that canonical policies tend to be too optimistic and over claim discoveries, we propose a novel policy that incorporates budget safety buffers. It turns out that a little more safety can greatly enhance efficiency — small additional logarithmic buffers suffice to reduce the regret from $\Omega(\sqrt{T})$ or even $\Omega(T)$ to $O(\ln^2 T)$. From a practical perspective, we extend the policy to the scenario with continuous arrival distributions, time-dependent information structures, as well as unknown T. We conduct both synthetic experiments and empirical applications on a time series data from New York City taxi passengers to validate the performance of our proposed policies. Our results emphasize how effective policies should be designed to reach a balance between circumventing wrong accept and reducing wrong reject in online resource allocation problems with exogenous budget replenishment.