ABSTRACT

As business experiments such as A/B or multivariate tests become more popular, firms seek to move beyond "test and learn" towards "earn while learning." Profitable experimentation, with the goal of continuous improvement, is framed as managing the multi-armed bandit problem. The contribution of the paper is to extend existing methods to handle various components of real-world marketing experiments (e.g., attributes of actions, unobserved differences in context, batched decisions). The benefits of the approach are demonstrated through a field experiment with an advertiser allocating ad impressions across many display ads on different websites to acquire customers. The approach is compared to benchmark methods via simulation studies.