

# 2020년 제 1회 통계세미나

고려대학교 통계연구소와 BK21 통계학교육연구팀이 다음과 같이 공동으로 세미나를 개최하오니 많은 참여 바랍니다.

일시 : 2020년 10월 30일 (금) 오전 11시

장소 : 고려대학교 정경관 201호

연사 : 임성빈 교수 (UNIST)

## **Optimal Algorithms for Stochastic Multi-Armed Bandits with Heavy Tailed Rewards**

<Abstract>

In this presentation, we consider stochastic multi-armed bandits (MABs) with heavy-tailed rewards, whose  $p$ -th moment is bounded by a constant for  $1 < p \leq 2$ . First, we propose a novel robust estimator where information about the constant is not required a priori, while other existing robust estimators demand the constant as prior information. Using this estimator, we propose a perturbation-based exploration strategy and develop a regret analysis scheme that provides upper and lower regret bounds for a general perturbation by revealing the relationship between the regret and the cumulative density function of the perturbation. From the proposed analysis scheme, we obtain gap-dependent and gap-independent upper and lower bounds of various perturbations. We also find the optimal hyperparameters for each perturbation, which can achieve the minimax optimal regret bound with respect to total rounds. In simulations, the proposed estimator shows favorable performance compared to existing robust estimators for various  $p$  values and, for MAB problems, the proposed perturbation strategy outperforms existing exploration methods. This work is accepted to NeurIPS 2020.

고려대학교 통계연구소  
BK21 통계학교육연구팀