2022년 제 10회 통계세미나

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

일시 : 2022년 9월 28일(수) 오후 5시
장소 : 고려대학교 정경관 501호
연사 : 박성오 교수 (성신여자대학교 통계학과)

Covariance Matrix Estimation with Incomplete Data and its Applications

<Abstract>

A sample covariance matrix of completely observed data is the key statistic in a large variety of multivariate statistical procedures, such as structured covariance/precision matrix estimation, principal component analysis, and testing of equality of mean vectors. However, when the data are partially observed, the sample covariance matrix from the available data is biased and does not provide valid multivariate procedures. To correct the bias, a simple adjustment method called inverse probability weighting (IPW) has been used in previous research, yielding the IPW estimator. The estimator can play the role of the sample covariance matrix in the missing data context, thus replacing it in off-the-shelf multivariate procedures such as the graphical lasso algorithm. However, theoretical properties (e.g. concentration) of the IPW estimator have been only established in earlier work under very simple missing structures; every variable of each sample is independently subject to missingness with equal probability.

In this talk, I will talk about the deviation inequality of the IPW estimator when observations are partially observed under general missing dependency. From it, we can derive the optimal convergence rate root-(log p / n) of the IPW estimator based on the element-wise maximum norm, even when two unrealistic assumptions (known mean and/or missing probabilities) frequently assumed to be known in the past work are relaxed. Finally, I will discuss three possible applications (graphical model, regularized linear regression, and testing cross-covariance) in statistics which the theoretical results are applicable to.

고려대학교 통계연구소
BK21 통계학교육연구팀