

愛媛大学データサイエンスセンター (CDSE) は、AI・統計解析・機械学習等の広義でのデータサイエンスと接点のある研究者、実務家、教育者を学内外から招聘し、講演していただくデータサイエンスセミナーを開催していきます。

第25回

愛媛大学

DS研究セミナー

2023.9.14(木) 16:30~18:00 参加費：無料
定員：オンライン300名
(Zoom・Youtube同時配信)

「空間データの特徴とサンプリングの空間的な偏りを踏まえた空間統計解析について」

標本が収集された地図上の位置情報を、緯度・経度、市区町村・大字小字などで含んでいるデータを空間データと呼んでいる。空間データの例として、賃貸アパートに関するデータにて、木造か鉄筋コンクリート造かといった建物の構造、築年数、階層、家賃といった賃貸アパートの属性の情報に加え、立地する住所が含まれているデータが挙げられる。このような空間データ、特に緯度・経度を含むものに対し、統計モデルをあてはめて分析を行う意義とその方法を紹介する。空間統計学において、分析の際に踏まえておく空間データの特徴として、空間的自己相関と空間的異質性が挙げられる。前者は距離の近い標本が似たような傾向を示すことを指し、例えば、属性が類似した賃貸アパートの家賃について、A地点とそれに隣接するB地点は似た価格であることが挙げられる。後者は統計モデルの観点から、モデルの誤差項の分散が空間的に不均一となることを指し、例えば、東京23区の賃貸アパートを例にとって考えると、田園調布、南麻布、南青山のような高級住宅街があちらこちらに点在し、家賃の空間的な均質性は満たされないことが挙げられる。これらの特徴に加え、空間データは観測対象となるエリアで満遍なく標本を収集できない状況がある。言い換えると、ある空間で標本が集中して取られる状況で、賃貸アパートの空間データであれば、住宅街では多いのに対し、農業地では少なくなる状況が挙げられる。先述した空間データの2点の特徴を踏まえつつ、空間的な偏りのあるデータの分析方法として、fused Lassoを用いた空間統計解析を紹介する。

講演者

山村 麻理子氏
(公益財団法人放射線影響研究所統計部研究員)



お申込み

右側Googleフォームもしくは下記URLよりお申込みください。
URL:<https://forms.gle/3n77Lk35rk5Hcybc8> 締切：9月12日(火)12:00
※資料は事前配布のみです。後日の配布は致しませんのでご了承ください。
※アーカイブ視聴はございません。当日の配信のみとなっております。

