

位置情報トラッキングデータを用いた主成分分析

～タクシードライバーの特性分析～

学籍番号 21711018

久保田ゼミ 4年 阿部 燎太郎

21711018ra@tama.ac.jp

タクシー業界の現状では運転手の確保が難しいといわれている。たとえ確保できたとしても入社早々の運転手は、お客様をのせて目的地に送るまでの道順を覚える必要があり、それらのスキルを習得しなければ熟練運転手並みにお客様を効率的に乗せることができないと考えられる。今のタクシー業界の新人ドライバーはベテランほど売り上げを稼げない。Uber アプリの登場で新人ドライバーが危機感を持っているために、稼げるようになるために研修をする必要がある。研修をするに当たって、ベテランドライバーの稼ぎ方を参考にしなければいけないので、ベテランドライバーの特徴を明らかにすることを研究の目的とした。

本研究では経営科学系研究部会連合協議会主催、令和元年度データ解析コンペティションで提供されたデータを用いることとした。そのデータには、タクシーの乗った所から降りた所までの 1km ごとの位置の記載があり、空車と実車などのステータスの情報がある。そこで、そのトラッキングを乗車して降車するまでのトリップのデータに加工した。そのトリップデータを用いて主成分分析の実施をした。

分析手法としては、R のパッケージである **radiant** を使用し、先述のとおり加工したデータを用いて主成分分析を実施した。変数としては、**dist0**、**1**、**2** (空車、実車、回送のトリップ平均距離)、**time0**、**1**、**2** (空車、実車、回送のトリップ平均時間)、**ttl** (全トリップにおける実車割合)、**worknight**(勤務時間における実車時間割合 (1 日))、**weekend**(勤務時間における週末労働割合)、**night**(勤務時間における深夜労働割合) 10 個の変数を用いた。固有値から 3 つの軸に絞ることとし、元のデータからトリップデータにまとめて、距離、時間、運賃を算出して、トリップデータから、ドライバーごと、日ごとに距離、時間、運賃を算出して分析した。

さらに、会社ごとに 3 つの属性の値から特徴を検討した。

まず、以下の 3 つの属性を判明させた。

- ・ 第 1 軸は長距離：長距離ドライバーが多い、実車時間距離が長いことが解る。
- ・ 第 2 軸は長実車時間：実車時間が長いことが解る。
- ・ 第 3 軸は週末・深夜：週末と深夜働いている。

分析結果等については、当日報告する。