

データサイエンス

keyword

- 統計的仮説検定
- 漸近理論
- データ解析



熊澤 吉起
Yoshiki Kumazawa

データサイエンス学部
教授

【専門分野】
・数理統計学

- 【プロフィール】**
- 1957 年 生まれ
 - 1983 年 6 月
大阪大学大学院
基礎工学研究科
数理系専攻
博士後期課程退学
 - 1983 年 7 月
滋賀大学経済学部
助手
 - 1984 年 10 月
同 講師
 - 1986 年 12 月
同 助教授
 - 1996 年 11 月
同 教授
 - 2017 年 4 月
データサイエンス学部
教授

【代表的な研究テーマ】

- 中途打ち切りデータに関する仮説検定
- データ解析

課題解決に役立つシーズの説明

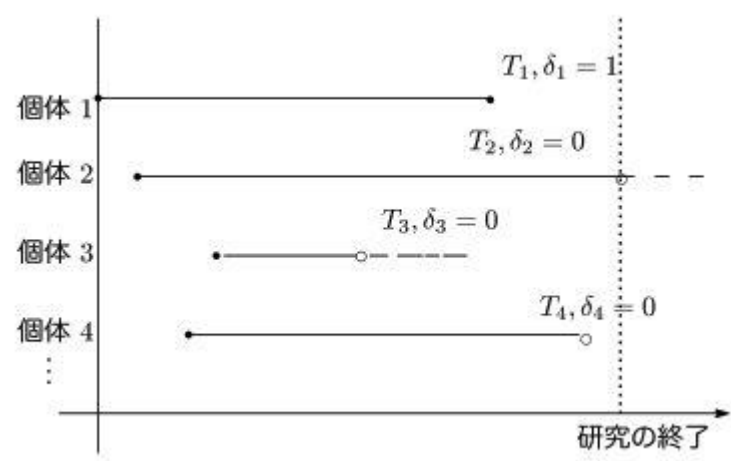
(1) 冒研究テーマについて

通常行われている検定問題は標本の観測の値がすべて最後まで観測されることを想定しています。研究しているテーマは信頼性・寿命等の検定問題において、途中で打ち切りを含む観測値がある場合にも適応した検定統計量の提案です。

信頼性・寿命検定において仮定された分布の多くは指数分布であり、これは故障率が一定であるという性質を持っています。しかし、初期故障、偶発故障、摩耗故障により、故障率が一定でなく、一般的に故障率はバスタブ曲線と言われるもので記述されます。これらの性質をもつ確率分布が検定問題の対立仮説の分布として考えられます。このような問題の性質から、検定問題はパラメトリックな検定問題でなく、ノンパラメトリックな検定問題として扱った方が応用の範囲が広がることになります。

工業製品の製品寿命や信頼性検査の際、時間内に検査が終了しないということがあり、また、費用の面から調査が途中で打ち切られるという状況が発生することがあります。工業製品でなく、病院の患者を考えると本来の調査項目でない事由により、退院、転院ということに対応します。このような状況においては、得られたデータは、打ち切られているかいないかの指標と、その時点までの時間になります。このデータの組に対していままで展開されてきた統計理論の枠組みでは解決できなく、多くの統計量は Kaplan-Meier 推定量によって表現されるものとなります。これは通常用いられる経験分布関数の拡張となっており、サンプルサイズを大きくしたときの大標本理論は、マルチンゲール理論に基づく点過程理論によって導かれます。

このような中途打ち切りデータの枠組みでの対立仮説に適した検定統計量を提案し、大標本での漸近分布を導き、漸近効率の良い検定統計量を求めています。



(2) 最近の関心事他

上記の検定問題の対立仮説としての確率分布の考察を行っています。また、オープンソースとして様々なプラットフォームで利用できる統計解析ソフトウェア R を用いて、シミュレーションをおこなっており、学部生・大学院生への指導の一環として、このソフトウェア R を用いた解説を pdf ファイルとして公開し、データマイニングやテキストマイニングで必要となるであろう解析手法について研究しています。