

小さな世界と大きな世界

渡辺澄夫

1 質問

質問

学生の皆様から次の質問をいただきました。

質問 「統計学を紹介する動画を見ていたら 小さな世界 と 大きな世界 という言葉が出てきたのですが、それらは何ですか」

このファイルでは上記の質問にお答えします。
答えのみ知りたい人は「4回答」をご覧ください。

小さな世界と大きな世界とは

いただいたご質問について、小さな世界と大きな世界は下記の動画で説明されています。

R. McElreath, Statistical Rethinking

<https://www.youtube.com/watch?v=4WVeICswXo4>

40:30 あたりで、small and large worlds が説明がされています。

言葉を作った人

小さな世界 という言葉は
主観ベイズ法の提案者 L.J. Savage に
よって作られたようです。

用語の混同に 注意してください。

このファイルで説明している small world は
主観ベイズ法 (or 決定理論) での用語であり、

ネットワーク解析における small world とは、
まったく関係はありません。

2 小さな世界

小さな世界

L. J. Savage は、主観ベイズ法を述べた本(1954)の中で Small world という言葉を、数学的に定義することは難しいが

跳ぶ前に 見ることができる

橋を渡る前に 渡れる橋はわかっている

という比喩で説明しています。

創始者による注意

Ken Binmore（数学者、経済学者）によると

「Savage は、主観ベイズ法の本(1954,p.16)で、主観ベイズ法は 小さな世界 のみに利用可能であると限定していた」

<https://link.springer.com/article/10.1007/s41412-017-0056-1>

(注) 正確には「ベイズ決定理論は小さな世界だけで意味がある」です。
主観ベイズ法は、ベイズ決定理論の特殊な利用例です。

小さな世界 ⇨ 主観ベイズ法が使える

Savage と Binmore によると

主観ベイズ法は小さな世界でだけ
利用してもよいものになります。

(注) 小さな世界において、統計的推測を行う方法はたくさんあります。主観ベイズ法は、その中のひとつです。小さな世界では、最尤法、一般のベイズ法、事後確率最大化法、などたくさんの方を用いることができます。主観ベイズ法は小さな世界でだけ利用することも可能です。

主観ベイズ法が使える前提

主観ベイズ法は、その定義から下記の場合にだけ使えます。

- (1) モデルは正しい。
- (2) 事前分布は主観で定められている。
- (3) サンプルは(1)(2)に従う確率変数。

主観ベイズ法が使える条件は定義から定まる

前ページの(1)(2)(3) のどれかが欠けると、
決定理論に基づくことができないので
主観ベイズ法を使うことはできません。
これは主観ベイズ法の定義から定まる性質です。

上記の意味がわからないかたは、下記を確認してください。

<http://watanabe-www.math.dis.titech.ac.jp/users/swatanab/decisiontheory.pdf>

小さな世界の例

小さな世界では 主観ベイズ法を使うこともできることから 小さな世界は モデルが明確に正しいとわかる世界であるということになります。例としては

さいころを振る

十分小さな有限集合の上で確率ゲームをする

などが小さな世界です。

主観ベイズ法は 小さな世界だけで

主観ベイズ法は、小さな世界だけを前提とし、モデルも事前分布も、データを得る前に正しいものが用意できることが必要です。データを得た後からのモデルのチェック、検定、比較、選択、統合、ハイパーパラメータの調節などはすべて許可されません。それらを行うと決定理論の前提が崩れ、主観ベイズ法ではなくなります。周辺尤度も交差検証も情報量規準も、それらを用いてモデルや事前分布を修正することはできません。

言葉だけの説明に注意しましょう。

「主観ベイズ法は真の確率分布がなくても使える」という説明は不適切な表現です。主観ベイズ法は、主観で決めた事前分布とモデルが真の分布を生成していると仮定します。それ以外の可能性を容認しません。

ある分野の研究者が「私の研究する〇〇学では確率分布では表されない現象を考察したいので主観ベイズ法を応用しよう」として進めると、その結果は、むしろ正反対で「真の確率分布を分析者が与えた解析」になります。

わかっていないこと

提案者 Savage による注意にも関わらず、主観ベイズ法が小さな世界でだけ利用可能なことが理解されていない理由はわかりません。無視してもよい些末なことでしょうか。

また Binmore の論文では「一貫性を犠牲にすることで決定理論の一般化を行う」とも書かれているのですが、このファイルの作成者は、それが何か(できるかどうか)は理解できていません。一貫性を放棄した決定理論とは何を意味しているのでしょうか。

3 大きな世界

大きな世界

Small world でないとき large world といいます。以下では、**大きな世界**と呼びます。あらかじめ用意したモデルでは記述できないことが起こりうるのが大きな世界の定義です。

◇ Binmore は、大きな世界の比喩として、局所的に見ると平坦であるが遠くまでみると実は曲がっている曲面をあげています。

◇ McElreath は、動画の中でコロンブスによる大陸発見を大きな世界の例にあげています。

大きな世界では

大きな世界ではモデルは正しくありません。

K.Binmore 「主観ベイズ法の一貫性は小さな世界でのみ利用可能である。大きな世界では一貫性は美德ではない」

K. Binmore, Rational Decisions. The Gorman Lectures in Economics, 2011.

<https://larspsyll.wordpress.com/2013/09/11/bayesian-inference-in-a-large-world/>

大きな世界では どうするのか

H. Gintis (経済学者)
(K. Benmore の本の書評の中で)

「大きな世界では【合理的決定とは何か】について Savage を超えた新しい定義が必要でしょう」

<http://www.umass.edu/preferen/gintis/BRDEJPubVer.pdf>

大きな世界を求めて

日本、統計数理研究所では1970年ころから現在に至るまで、「データが増えるにつれて詳細な状況がわかってくる」という設定における統計学が深く広く建設されてきました。

すなわち、情報量統計学は、研究の最初から、大きな世界、未知の世界を対象としていたことになります。

大きな世界のベイズ法とは

実世界は大きな世界です。大きな世界のベイズ法の現代的な紹介である「Gelman, et. al., BDA3」や「McElreath, Statistical Rethinking」では

「ベイズ法では、データとパラメータの同時分布をモデルとする。データとパラメータを区別せず、観測できないものは分布のままとし、知りたいものの条件つき確率を計算する」

という**一般のベイズ法**を推奨しています。

大きな世界を考える基盤は

大きな世界では、一般のベイズ法が利用されています。すなわち、モデルと事前分布の組がモデリングされるものです。大きな世界のベイズ法では、定義から決定理論も主観ベイズ法も基礎とすることはできません。

一般のベイズ法においては決定理論は使えませんが、決定理論で最小化しようとしていた量(例:汎化誤差)を推測できる理論を作ることは可能であり、それが大きな世界の統計学の基盤になります。

どちらの世界も研究の対象である。

研究においては、様々な条件のもとで、いろいろな問題が考察され、それら全体が学問となっていきます。

統計学を整備していくとき、モデルが正しいという前提のもとで、種々の統計量の性質を調べることには、十分な意義があります(例:検定の理論)。

研究においては、常に大きな世界のみを考えなくてはならないということはありません。

大きな世界は初めての人には難しい

初めて統計学を習う人に、大きな世界の話の説明することはたいへん難しいことであると思います。モデルを作るにも関わらず、そのモデルが正しくないことが出発点となるからです。

初めての人には小さな世界の話から始めるということも可能ですが、小さな世界だけに限定されている方法(主観ベイズ法)を一般のベイズ法の基礎であるかのように説明することは適切でないと思います。Binmore によれば、Savage の注意にも関わらず主観ベイズ法(と決定理論)が大きな世界の基礎でもあると誤解している(あるいは事実には気づかない)学者が少なくないとのことです。

統計学の重要ポイントは

大きな世界の統計学において最も重要な点は、

「モデリングする」

ということを理解する点だと思います。大きな世界では、モデリングが正しい保証はありませんが、モデリングなしでは何もできません。モデリングがどのくらい間違っているかを知ることが必要です。(ベイズか最尤かを問わず)ここが統計学を理解するのに欠かすことができない大切なポイントであると思います。

初めての人に統計学を紹介する担当者のかたには、ぜひ、ご自身でお考えいただき、ご自身の言葉で説明していただきたい点です。

4 回答

質問への回答

小さな世界は、「モデルが正しいことが明確で、主観ベイズ法を用いることも可能なケース」を表す言葉です。

大きな世界は、「モデルは候補にすぎず、一般のベイズ法が必要になるケース」を表す言葉です。

5 参考

(参考) 情報量統計学は大きな世界で

赤池弘次先生 「現実の世界では、データが増えるにつれて少しずつ詳細な構造が現れてきます。AICによるモデル選択は、そのときそのときのサンプルサイズにおいて、統計的揺らぎとバイアスのぎりぎりのバランスを取ることで、最小の予測誤差を平均的に与えるモデルを選び出します」

↑ 時系列モデルにおける証明も柴田里程先生によりなされています。

↑ すなわち情報量統計学では当初から大きな世界が想定されていました。数理科学にもとづいて考えることがいかに大切かがわかる例です。

(参考) 箱庭から出た人工知能

1985年ころ、「人工知能は、人為的に作られた箱庭の中では良好に動作するが、無限の未知が生じる現実の世界 open world で使うことはできないだろう」と考えられていました。

今日の人工知能が、まだ発展の途中とはいえ、画像認識、音声認識、機械翻訳で実務に使われるようになったのは、データの量、コンピュータの容量と速さ、ネットワーク、そして深層学習が、箱庭の外の世界へ踏み出し始めたからのようです。

機械学習が人工知能の実現を目標として持つならば、大きな世界こそが研究の対象とするものであることになります。

(参考) 主観と客観 現代の視点

Gelman と Hennig は、統計的推測に恣意性が大きな影響を及ぼすのは事前分布よりも、「何を分析するか」や「統計モデルの構成」においてであり、そうした広い意味での主観なしに統計的推測はできないと述べています。そのことは前提としたうえで、モデルのチェック・比較・検定・選択等を行ってできるだけ客観的であろうとすることを「**透明性・合意性・公平性の美徳**」と呼んでいます。またそのようにしないことを「**多様な視点を与える美徳**」と呼び、【主観、客観】という古くて対立的な用語は不毛なので使用をやめましょうと提案しています。

<https://rss.onlinelibrary.wiley.com/doi/full/10.1111/rssa.12276>