

2015年
12月1日
火曜日

白井 洸志 専任講師 (ミクロ経済学)

The Law of Decreasing Credibility

——データなる信用の法則——

昨今では誰もが様々なデータにアクセスし、それを解析することが可能である。結果として、我々の周囲には「データによると」を枕にした主張が氾濫している。しかしながら、データそのものからセンセーショナルな結論が得られることは稀であり、多くの場合、それらは暗黙のうちに課された制約的な仮定の下で得られている。さらに言えば、同一のデータを異なる仮定の下で分析すれば全く異なる分析結果が得られることがむしろ普通である。以下では、C.F. Manski (著) Identification for prediction and decision の第10章を基に、この点について具体例を用いて述べる。

いま、貧困層子女に対する早期教育(3歳児4歳児教育)が彼らの高校卒業率に及ぼす影響を測定するため、以下のような社会実験を行ったとする。ある街の貧困層子女を無作為に二集団抽出し、グループ1については何も行わず、グループ2については早期教育を施す。追跡調査により19歳時点での高校卒業率をグループ毎に測定し、

- ・グループ1…49%
- ・グループ2…67%

というデータが得られたとしよう。各グループは無作為抽出されているので、

- ・現状の高校卒業率 49%
- ・仮に貧困層子女全員が早期教育を受講した場合の高校卒業率 67%

と考えてよい。

ここで次のような問題を考える。

Q. 早期教育施設を設置し無料開放したならば、貧困層の高校卒業率ほどの程度になるか?

この問に対して、データをフル活用して得られる解答は「16% (100%)」である。解は一意に定まらないどころか、現状より改善するか否かすらデータは語らない。紙幅の関係で証明を与えることはできないが、いくつかヒントを挙げておく。まず人々は

1. 早期教育を受けても受けなくて
2. 早期教育を受けた場合にのみ高校卒業するタイプ
3. 早期教育を受けた場合にのみ高校卒業しないタイプ
4. 早期教育を受けても受けなくても高校卒業しないタイプ

のいずれかに必ず分類される。また早期教育施設が開放されたとしても、それを利用するか否かは自由意思であるという点にも注意されたい。データから全員が早期教育を受講した場合の高校卒業率と誰も早期教育を受講しない場合の高校卒業率はわかっているため、各タイプがどんな比率で混在しているのか部分的に導くことが出来るだろう。

上で述べた予測値を改善する唯一の方法は、人々の行動形態や多様性に予め仮定を置いて分析することである。例えば、

仮定1. 早期教育が悪影響を及ぼすことはない(タイプ3は存在しない)。

として分析すれば、予測値は49% (67%となる。また、

仮定2. 各タイプは自らがより高校卒業できるように早期教育受講の有無を決定する。

として分析すると、予測値は67% (100%となる。少々テクニカルな仮定であるが、

仮定3. 早期教育が高校卒業率に及ぼす効果は統計的に独立である。とすれば、予測値は33% (83%となる。さらに、仮定2と仮定3を両方用いると、ズバリ83%が予測値として導出される。

上のような、「データから言えることは極めて弱く、強い結論を得るためには制約的な仮定を要する。また、仮定次第で結論は異なる。」という性質は実証分析一般に成り立つ基本原理である。「データによると」を枕とする主張について、それがどのような仮定に依って立っているのか、またその仮定が妥当なのか否か、それらを見極める力も経済学部で身につけるべき能力の一つである。