

項目応答理論概要 【評価版】

本 whitepaper では項目応答理論 (IRT: Item response theory) の機能についてその概要を紹介します。

1. 項目応答理論とは

Example 1

Example 2

補足 1

1. 項目応答理論とは

研究者はときに能力 (abilities) や性格特性 (personality traits) といった観測不能な資質について調査を行いたくなることがあります。[IRT] マニュアルにおいてはこれら観測不能な資質について言及する際に潜在特性 (latent trait) という用語が用いられていますが、能力という用語がそのまま使われていることもあるので注意してください。

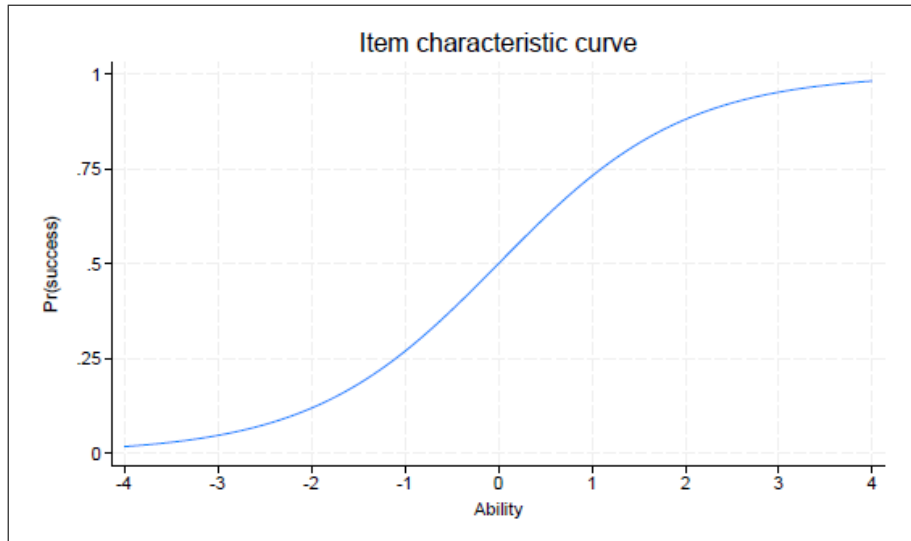
潜在特性は観測不能であるため直接計測することはできませんが、instrument を用いて定量化することができます。Instrument というのは潜在特性のレベルを計測するという目的で企画された項目 (items) の集合体のことを意味します。数学的能力 (潜在特性) を計測しようとする研究者が 100 個の設問 (items) からなるテスト (instrument) を企画するといったケースがその典型例です。

Instrument を企画する場合、あるいは instrument を介して得られたデータを分析しようとする場合、個々の項目、あるいは一群の項目が、対象となる特性とどのような関わりを持つのが、研究者にとっての関心事となります。IRT モデルを用いるとそれらの関係について探究することが可能になります。

IRT モデルは認知や性格といった特性の研究に幅広く使用されているわけですが、その応用事例としては次のようなものをあげることができます。

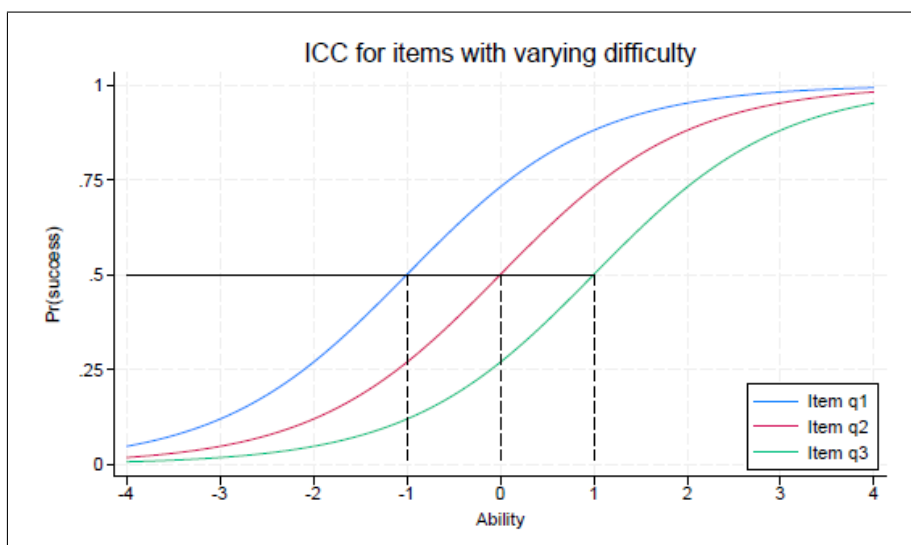
- 小学児童における computer anxiety の計測 (King and Bond 1996)
- HIV に感染した成人を対象とした身体機能の評価 (Wu, Hays, Kelly, Malitz, and Bozzette 1997)
- 栄養士による公共政策への関与の度合の計測 (Boardley, Fox, and Robinson 1999)

IRT において中心的役割を果たすものとしては項目特性曲線 (ICC: item characteristic curve) をあげることができます。ICC はある人が与えられた項目 (個々の設問) に対して“成功”する確率を表現したものです。次のグラフ中には能力の計測を企図した項目に対する ICC が示されています。能力の増加に伴い成功する確率も高くなる点に注意してください。



ICC は項目ごとに異なったものとなります。ある項目に対して成功する確率は潜在特性のレベルと当該項目の属性双方の関数と言えます。潜在特性は一般に θ と表記されます。調査対象者に対する θ の値は person location と呼ばれます。一方、項目の属性は難易度 (difficulty) と弁別力 (discrimination) と呼ばれるパラメータであり、IRT モデルの中で推定されることになります。

難易度パラメータ — item location とも呼ばれ、通常 b と表記される — は能力スケール上における項目の位置を表します。例えば次のグラフはそれぞれ $-1, 0, 1$ という値の難易度パラメータを持つ項目 q1, q2, q3 に対する ICC をプロットしたものです。



項目 q1 は最も難易度が低く、q3 は最も難易度が高いものと言えます。難易度の変化によって ICC が能力スケール（水平軸、 x 軸）に沿ってシフトする点に注意してください。どの能力レベルにおいても項目 q1 に対する成功確率は q2, q3 に対するものより高くなっています。項目 q1 に対して成功が期待されるには能力レベルが -1 よりも高ければ良いことになるので、項目 q1 は q2, q3 よりもやさしいと言えるわけです。これに対し、項目 q2 に対して成功が期待できるためには能力レベルが 0 よりも高いこと、項目 q3 に対して成功が期待できるためには能力レベルが 1 よりも高いことが要求されます。

潜在特性のすべてのレベル間で識別が行える instrument を企画しようと思ったら、すべての特性レベルをカバーする難易度を持った項目を用意する必要があります。

2 番目の項目パラメータである弁別力 — 通常 a と表記される — は ICC の勾配に関係したものです。この項目パラメータは、その項目の難易度近辺において能力を増加させたときに、どれだけ早く成功確率が変化するかを教えてください。弁別力の値の大きな項目は潜在特性と当該項目に関する成功確率との間に強い相関を持っています。言い換えるなら、大きな弁別力を有する項目は潜在特性の低いレベルと高いレベルとをより良く識別できるものと言えます。

評価版では割愛しています。

▷ Example 1: 二項 IRT モデル

評価版では割愛しています。

▷ Example 2: カテゴリカル IRT モデル

評価版では割愛しています。

補足 1 – collapse 操作

評価版では割愛しています。

■