

項目反応理論

2021 年 7 月 13 日

4 項目母数の推定

4.2 周辺最尤推定法

同時最尤推定法では、最大化する尤度関数が多くなってしまうという欠点がある。最大化する尤度関数の中から、被験者母数をなくすことができれば、最大化に関する母数を大幅に減らすことができる。また被験者を追加してデータを増やしても未知数が増えることはないので推定値が安定するというメリットもある。つまり、反応パターンが与えられたときの項目母数だけの尤度関数を構成すればよい。この考えを周辺最尤推定法と呼ぶ。 n 個の項目に対する被験者 i の反応パターンベクトル \mathbf{u}_i が観察される確率は $f(\mathbf{u}_i | \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{c})$ であった。目標はここから被験者母数 θ_i を消すことである。

4.2.1 周辺化

同じ形の赤い壺、青い壺、緑の壺に基石がそれぞれ 10 個入っておりそのうち黒い基石がそれぞれ 3 個、4 個、5 個入っている。目を閉じたまま基石をとる思考を考える。ここで選んだ壺から黒い石が取り出される確率は、壺の色が与えられたときに決定されるから確率表記は $p(\text{黒} | \text{壺の色})$ であり、「壺の色」は母数である。式にすると、

$$0.3 = p(\text{黒} | \text{赤}) \quad (4.15)$$

$$0.4 = p(\text{黒} | \text{青}) \quad (4.16)$$

$$0.5 = p(\text{黒} | \text{緑}) \quad (4.17)$$

である。ここで壺の色は考えないことにする。つまり、壺の色という母数を消去して、黒の石が取り出される確率を考える。このとき、考える対象ではない母数を局外母数という。これを考えるためには、「壺の色」自身の確率分布 $p(\text{壺の色})$ を利用する。

例えば、赤い壺、青い壺、緑の壺が選ばれる確率が問う確率ならば、黒い石が取り出される確率は、

$$p(\text{黒}) = \frac{1}{3}(p(\text{黒} | \text{赤}) + p(\text{黒} | \text{青}) + p(\text{黒} | \text{緑})) = \frac{1}{4} \quad (4.18)$$

であり、これによって壺の色を考慮せず考えることができる。ここで、消去される母数の分布が一様分布である必要はない。例えば、赤い壺、青い壺、緑の壺が選ばれる確率がそれぞれ、0.5, 0.3, 0.2 であるとする、黒い碁石が取り出される確率は、 $p(\text{黒}) = p(\text{赤い壺を選ぶ確率})p(\text{黒}|\text{赤}) + p(\text{青の壺を選ぶ確率})p(\text{黒}|\text{青}) + p(\text{緑の壺を選ぶ確率})p(\text{黒}|\text{緑})$ で考えることができ式にすると、

$$p(\text{黒}) = p(\text{赤})p(\text{黒}|\text{赤}) + p(\text{青})p(\text{黒}|\text{青}) + p(\text{緑})p(\text{黒}|\text{緑}) = 0.37 \quad (4.19)$$

である。局外母数を確率分布から求めるためには、局外確率自身の分布と局外母数で条件つけられた確率分布との積和を計算すればよい。これを周変化による局外母数の消去という。

4.2.2 周辺尤度関数

本節では、被験者母数を局外母数として確率分布から、 θ を消去する。 θ 自身の分布を $g(\theta)$ とする。

ここでは、説明変数は離散型として4つの値 $\theta = 1.0, 0.5, -0.5, -1.0$ しかとらないと仮定する。例えば、それらを「優」、「良」、「可」、「不可」とみなしそのような学生が大学の中に0.3, 0.2, 0.2, 0.3 の確率で存在しているものとする。このとき、母集団から被験者を選び出し、受験させある反応パターン \mathbf{u}_i を得る。 $f(\mathbf{u}|\theta, \mathbf{a}, \mathbf{b}, \mathbf{c})$ の被験者母数に具体的な尺度値を代入して、 $g(\theta)$ との積和を求める。式は、

$$f(\mathbf{u}|\theta, \mathbf{a}, \mathbf{b}, \mathbf{c}) = g(1.0)f(\mathbf{u}_i|1.0, \mathbf{a}, \mathbf{b}, \mathbf{c}) + g(0.5)f(\mathbf{u}_i|0.5, \mathbf{a}, \mathbf{b}, \mathbf{c}) \quad (1)$$

$$+ g(-0.5)f(\mathbf{u}_i|-0.5, \mathbf{a}, \mathbf{b}, \mathbf{c}) + g(-1.0)f(\mathbf{u}_i|-1.0, \mathbf{a}, \mathbf{b}, \mathbf{c}) \quad (2)$$

$$= 0.3f(\mathbf{u}_i|1.0, \mathbf{a}, \mathbf{b}, \mathbf{c}) + 0.2f(\mathbf{u}_i|0.5, \mathbf{a}, \mathbf{b}, \mathbf{c}) \quad (3)$$

$$+ 0.2f(\mathbf{u}_i|-0.5, \mathbf{a}, \mathbf{b}, \mathbf{c}) + 0.3f(\mathbf{u}_i|-1.0, \mathbf{a}, \mathbf{b}, \mathbf{c}) \quad (4)$$

のように導く。周変化によって被験者母数を消去することができた。

ここでは、離散的な説明変数を考えたが連続型はどのように考えるのだろうか。これは、階級の幅を狭めることができれば実質離散型として扱うことができる。例えば、小学1年生の体重を記録する際に、10kg から 35kg までを 0.2kg 刻みの離散変数として扱うことができる。そこで、 K 個の値をとる階級地 θ_k の分布 $g(\theta_k)$ を考える。これを、多項分布という。添え字 k は被験者ではなくヒストグラムの階級値についていることに注意。このとき、母集団から1人の被験者を抽出し受験させ \mathbf{u}_i という結果を得る確率は、

$$f(\mathbf{u}_i|\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{k=1}^K g(\theta_k)f(\mathbf{u}_i|\theta_k, \mathbf{a}, \mathbf{b}, \mathbf{c}) \quad (4.20)$$

と表現され、被験者母数は消去される。ただしこの場合は被験者母数の母集団におけるヒストグラムが既知であることが必要である。

先に述べたように、連続変数を細かい階級幅で近似することは問題ない。しかし、実際にはその分布が既知であることは期待できないことが多くそれでは、式 4.20 は利用できない。被験

者母数を本来のモデルとおりに連続変数として扱うためにその分布に標準正規分布、

$$g(\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\theta^2\right) \quad (4.21)$$

を仮定してみる。