

項目反応理論

2021 年 10 月 13 日

4.3 ベイズ推定法

本節では 3.4 節で導入したベイズ推定法を項目母数の推定に利用する方法を論じていく。周辺推定法では、被験者が全員正答、全員誤答した項目の項目母数は推定することができなかったが、ベイズ推定法なら可能である。それ自体が問題になるわけではないが、事前分布が適切である場合にはベイズ推定法は実用的には便利な方法である。ここでベイズの公式を再掲したものが以下である。

$$f(A|B) = \frac{f(A)f(B|A)}{f(B)} \quad (4.29)$$

ここでは、反応パターンが与えられたときの母数に事後分布を求めるので、 A は θ, a, b, c に相当し、 B は、 U に相当する。

4.3.1 同時尤度による方法

はじめに、 $f(A)$ に相当する事前分布を導く。ここでは、全ての母数は独立しており、ある確率分布 $g(\cdot)$ に従ってランダムに分布していることを仮定している。独立の仮定の下で、複数の母数の事前分布はその総積で

$$g(\theta, a, b, c) = \prod_{i=1}^N g(\theta_i) \prod_{j=1}^n g(a_j)g(b_j)g(c_j) \quad (4.30)$$

と表現される。これを同時事前分布という。被験者母数の事前分布 $g(\theta_i)$ には正規分布が識別力の事前分布 $g(a_j)$ には、対数正規分布が、困難度の事前分布 $g(b_j)$ には正規分布が、当て推量の事前分布 $g(c_j)$ にはベータ分布が仮定されることが多い。式 (4.9) の $L(U|\theta, a, b, c) = f(U|\theta, a, b, c)$ と式 (4.29) $f(A|B) = \frac{f(A)f(B|A)}{f(B)}$ を用いると、

$$f(L|\theta, a, b, c) = \frac{g(\theta, a, b, c)L(U|\theta, a, b, c)}{f(U)} \quad (4.31)$$

のように表現することができる。この式を θ, a, b, c に関して最大化し、ベイズ推定値を得る。実際には積の連なりであることが多いので対数変換を行って母数で偏微分し 0 と置き連立方程式を解く。母数で偏微分した時点で、分母の $f(U)$ は 0 になるので、分子の最大化を考えればよい。

4.3.2 周辺尤度による方法

ここでは、周辺化によって被験者母数を消去して母数の数を減らす方法を紹介する。まず、周辺化により被験者母数を消去すると、項目母数の同時事後分布は

$$f(\mathbf{a}, \mathbf{b}, \mathbf{c} | \mathbf{U}) = \int f(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{c} | \mathbf{U}) d\boldsymbol{\theta} \quad (4.32)$$

と表現される。(理由は以下で説明する)

連続型確率変数

連続型確率変数 \mathbf{X}, \mathbf{Y} では、それぞれの実現値 x, y と確率密度 $f(x), f(y)$ に関して以下が成り立つ。

$$\int f(x|y) dx = 1 \quad (4.33)$$

から確率変数の連なりに置き換えて

$$\int f(x, y|z) dy = f(x|z) \quad (4.34)$$

式 (4.32) を $\mathbf{a}, \mathbf{b}, \mathbf{c}$ に関して最大化し、ベイズ推定値を得る。

4.3.3 被験者母数の事前分布

上記に示したように、被験者母数の事前分布には正規分布が仮定されることが多い。正規分布の母数は平均 μ_{θ} と分散 σ_{θ}^2 である。

4.3.4 困難度の事前分布

困難度の事前分布にも正規分布が仮定されることが多い。被験者母数の事前分布を $\mu_{\theta} = 0, \sigma_{\theta}^2 = 1$ で設定する場合にはそれより広い範囲をカバーする事前分布を仮定する必要がある。あるソフトウェアでは $\mu_{\theta} = 0, \sigma_{\theta}^2 = 4$ で設定されているものもある。

4.3.5 識別力の事前分布

識別力は原理的には、負の値もとる得る。そこで、識別力の事前分布には正の範囲で定義される対数正規分布が仮定されることが多い。

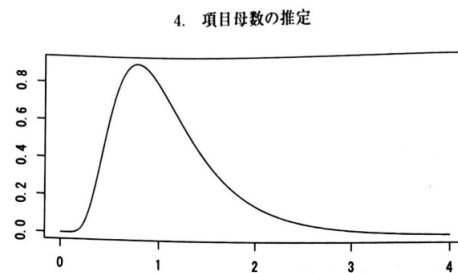


図 4.3: 識別力の事前分布 (対数正規分布)

4.3.6 当て推量の事前分布

当て推量は確率を表現する母数であるため、事前分布も $(0, 1)$ 区間で定義されることが望ましい。このため事前分布はベータ分布が仮定されることが多い。ベータ分布の母数は α, β であり、ベータ分布の最頻値は、

$$\frac{\alpha - 1}{\alpha + \beta - 2} \quad (4.35)$$

であることが知られている。当て推量の尤もらしい目目安は選択肢数の逆数 c と考えられる。目安 c を事前分布の最頻値の一致させるならば、例えば、

$$\alpha = 20 \times c + 1 \quad (1)$$

$$\beta = 20 \times (1 - c) + 1 \quad (4.36)$$

とする。5肢選択ならば $c = 0.2$ とし $\alpha = 5, \beta = 17$ のベータ分布を考える。

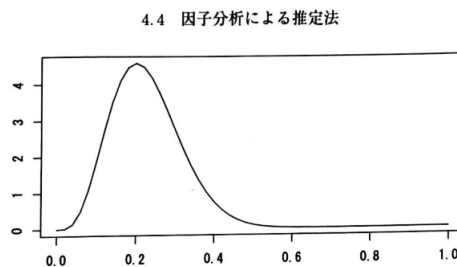


図 4.4: 当て推量の事前分布 (ベータ分布)

上記の図を見てみると、分布の最頻値が 0.2 であることが確認できる。