

## 自由エネルギー原理から道德教育へ

—— 自閉症をめぐる ——

竹 中 均

### 一 自由エネルギー原理とベイズ推論

自閉症は現在、脳の機能障害によって生じると想定されている。ゆえに、社会性の障害を特徴とする自閉症を考察するには、脳の働きについて考える必要がある。ところで近年、脳の働きを説明するための新しい理論として、「自由エネルギー原理」が注目されている。この原理とそこから導き出される様々な計算理論は広汎な射程を持つが、その一つとして統合失調症や自閉症のメカニズムを説明しようという試みがある。自由エネルギー原理の立場によれば自閉症はどのように説明されるのか。そしてその説明が、自閉症者とその生き方に対する人文・社会科学的アプローチにどんなインパクトを与えるのか。その可能性について論じたい。まず、「自由エネルギー」という、人文社会系にとっては馴染みの少ない用語が理論の中核をなしている点が印象的である。自由エネルギーは熱力学や（エントロピー概念を駆使する）情報理論において重要な概念であるが、それが自閉症の解明とどう関わるのだろうか。

二〇二二年に刊行・翻訳された『能動的推論—心、脳、行動の自由エネルギー原理』は、自由エネルギー原理の入門書である（パー他 二〇二二：三二九）。それによれば自由エネルギー原理の中核である「能動的推論」を説明するには「王道」と「常道」がある（パー他 二〇二二：七）。王道はトップダウンに、自由エネルギーから出発する説明であり、常道はボトムアップに、ベイズの定理から出発する説明である（パー他 二〇二二：八）。常道の方は「ベイズ脳という概念から始まる」（パー他 二〇二二：七）。

つまり、理論の形成過程としてはベイズ脳仮説が出发点であり、そこから、〈脳は自由エネルギーの最小化を目指す〉という原理が提唱された。その上で方向を逆転させて、狭義の自由エネルギー原理を出发点にして、ベイズ脳のみならず、広汎な生命活動を統一的に説明しようとする大胆な試みが、王道としての自由エネルギー原理である。

そこで当面、自由エネルギーは後回しにして、ベイズ脳仮説から話を始める。例えば知覚過程では、脳は原理的に、外界という環境について直接的に知り得ない。脳に出来るのは、環境についての「生成モデル」を自らの内側に作り出すこと、そして、五感を通じて感覚信号を観測することの二つである。脳はこれら二つを利用して、環境の状態（原理的に知り得ないために「隠れ

状態」や「隠れ原因」と呼ばれる)を推論せざるを得ない。そのための道具立てがベイズ推論だという仮説である(乾・阪口 二〇二一:二五)。

ベイズ推論ではまず「事前確率」 $p(x)$ を設定する。もっとも、確率とは言っても、脳が扱う推論の大部分は、〈壺の中に白と赤の玉がいくつか入っていて、その中からランダムに一個を取り出す〉というようなモデルでは対応出来ない。有限の離散的な数ではなく、無限の連続量を扱わなくてはならない。例えば明日の天気の可能性は実際のところ、晴れか雨かの二項対立のような離散的な数ではないので、無限の連続量である。晴れと雨との間には、無限のスペクトラムがある。したがって本来は、確率ではなく確率分布すなわち確率密度関数が扱われるべきである(豊田 二〇一六:七)。しかしながら本稿では簡略化して、「確率」と表記して論じることにする。

事前確率は、脳が感覚信号を観測する以前に作り出した、特定の「隠れ状態」が環境に生じる確率である。もちろん、感覚信号というデータ無しの状態での推論なので、その確率が正しいかどうかは分からない。(乾・阪口 二〇二一:二五)。そのため事前確率は「事前信念」とも呼ばれる。

もう一つの確率は、「尤度」と呼ばれる「条件付き確率」 $p(y|x)$ である。隠れ状態  $x$  が生じているという条件の下で感覚信号  $y$  が得られるだろう確率である。例えば、外界にリンゴがあるという隠れ状態が実際に存在する場合、赤くて丸いイメージが視覚に入ってくるという事象の生起がどの程度もっとも(尤も)らしいかを示す確率である。外界に本当にリンゴが存在すれば、脳が感覚器官を通じて赤くて丸い視覚イメージを得る可能性はおそらく高いだろうと予測される。だが、脳が外界を原理的に把握出来ない以上、リンゴの存在と赤くて丸い視覚イメージとの間の因果関係については確率的にしか判断出来ず、十分に客観的とは言えない。それはまさに尤もらしさであり、正しさではない。

以上のようにして、事前確率と尤度の両方が与えられれば、そこから計算して、感覚信号  $y$  が観測されたという条件の下で、外界に隠れ状態  $x$  が存在しているだろう確率、すなわち「事後確率」 $p(x|y)$ が(ストレートにはではないが)計算出来ると考える。後述するように、これがベイズ推論である。事後確率は事前確率とは違い、感覚信号というデータを考慮した後の確率なので「事後」と呼ばれるが、感覚信号を入手してもなお外界の真実は原理的に分かり得ないので、「事後信念」とも言う。例えば、丸くて赤いイメージが見える際に、その原因が外界に存在するリンゴである確率がそれである。

しかし、その確率の数値を実際に計算するためには、そもそも、リンゴ以外の原因の可能性全てを把握する必要がある。なぜなら確率とは、当該の事象の可能性を全ての可能性で割り算したものだからだ。だが、〈壺の中に白と赤の玉がいくつか入っていて、その中からランダムに一個を取り出す〉というようなモデルならばともかく、脳が現実直面している状況では、全ての可能性を網羅することは実質的に不可能である以上、この計算は困難である。

ともあれここで気付くのは、尤度は〈外界のリングから丸くて赤い感覚信号へ〉というのに対し、事後確率は〈丸くて赤い感覚信号から外界のリングへ〉という逆方向だという点である。一見すると奇妙なこの計算はどのようにして可能なのか。それを説明するのがベイズの定理である。

事前確率  $p(x)$  と尤度  $p(y|x)$  を掛け算すると、隠れ状態  $x$  について知りうる全て（隠れ状態  $x$  がどんな確率で存在し、そこから感覚入力  $y$  がどんな確率で生じるか）を把握したことになる。それは、外界環境について脳が作り出した「生成モデル」つまり同時確率  $p(y, x)$  である。この記号では、尤度において「|」（given）で表現されるような条件付けは存在せず、無条件に感覚信号  $y$  と隠れ状態  $x$  の確率が示されている（乾・阪口 二〇二一：二六）。

「生成モデル」を導き出すには別のやり方もある。それは、感覚信号の確率  $p(y)$ （新出、「周辺尤度」とも呼ばれる）と事後確率  $p(x|y)$ （既出）の掛け算だ。この掛け算の結果もやはり、同時確率  $p(x, y)$  すなわち生成モデルである。このようにして、以下の二つの等式が導き出された。

$$p(x, y) = p(y|x)p(x)$$

$$p(x, y) = p(x|y)p(y)$$

トーマス・パーらの表現によれば「ベイズ規則は、2つの事象の同時確率が、2番目の事象が与えられたときの1番目の事象の確率に2番目の事象の確率を掛けたものであり、これは1番目の事象が与えられたときの2番目の事象の確率に、1番目の事象の確率を掛けたものと等しいということに他ならない」（パー他 二〇二二：一七）。

これら二つの等式を合体すれば、次の等式が得られる。

$$p(x, y) = p(x|y)p(y) = p(y|x)p(x)$$

さらにこれから次の等式が得られる。

$$p(x|y) = p(y|x)p(x) \div p(y)$$

つまり、事前確率  $p(x)$  と尤度  $p(y|x)$  から、事後確率  $p(x|y)$  が計算出来る。ただし、そのために本当は周辺尤度  $p(y)$  が既知でなければならない。だが既述のように、周辺尤度を知ることは難しい。なぜなら、特定の感覚信号  $y$  が生じる確率を計算するためには、 $y$  以外のあらゆる感覚信号の可能性を全て把握しなければならないからである。たとえ話で言えば、壺の中にどんな色の玉が何個入っているか、そもそも玉以外の物体が入っていないかどうか予め分からなければ、その壺から赤い玉が取り出される確率を計算出来ないのと同じである。

とは言うものの、式の名分母が分からなくても、分子（事前確率  $p(x)$  とそれに対応する尤度  $p(y|x)$ ）さえ分かれば、異なる分子ごとの計算結果つまり事後確率の大小だけは相互比較出来る。結果的に、いくつもの可能性のうち、どれが事後確率最大かは判断出来る。推論において実際に必要なのは、どれが最大かを知ることであって、個々の事後確率の数値自体は知らなくても構わない場合が多い。このように、事後確率  $p(x|y)$  が最大になる隠れ状態  $x$  を求める推定を「最大事後確率推定」と呼ぶ（乾・阪口 二〇二一：二七、小島 二〇一五：一八六）。

ここで注目すべきは、この推論が時間の流れの中で展開する点である。つまり、感覚信号の観測は一回とは限らず、実際の脳がそうであるように、何度も繰り返される。だが、何度繰り返しても、脳が環境の現実を真に把握することは原理的にあり得ない。それでもなお観測結果は蓄積されていき、脳が行う推論に変化を与え続ける。つまり、脳が抱く信念と呼ぶ他ない事前確率  $p(x)$  を出発点として、最初 sensation 信号  $y_1$  を考慮することによってもたらされた事後確率  $p(x|y_1)$  は、次の sensation 信号  $y_2$  が到来した後、新たな二巡目の事前確率つまり新たな出発点として再帰的に導入され、時間の中でベイズ推論の流れを作動させる。

だからと言って、二巡目の計算の際、最初 sensation 信号  $y_1$  は消え去ったのではない。ただ、最初事前確率とは違う数値となった二巡目の事前確率の中に、見えない形で繰り返されている。たとえ話で言えば、二巡目の事前確率は、最初 sensation 信号  $y_1$  を無意識的に記憶している。

もちろん、これで推論が終わるわけではない。感覚信号の観測が繰り返される度に、事後確率が次の事前確率になるという再帰の流れはいつまでも続く。このような状況は「ベイズ更新」と呼ばれる。

## 二 ベイズ更新と逐次的合理性

古典統計ならば、感覚信号の観測データが全て揃ってからそれらを一体化し、一覧可能な形にまとめた上で初めて計算過程へ入っていくだろう。その結果として算出される数値は、ベイズ推論のような確率ではなく、定まった値となる。もちろん、定まったとは言っても有意水準が考慮されるが、古典統計の推論自体は、ベイズ統計と比較すると確率的性格が強くない。例えば、有意差の検定では、有意差があるかないかが二項対立的に判断される。敢えて言えば、外界という環境に確固とした真実があり、推論によってその真実へ接近していくという前提に立っているように思える。たとえ話で言えば、野球において審判が間違える可能性は常に考慮されているものの、判定はストライクかボールかの二項対立で下されるのと似ている。

それに比べ、再帰的にベイズ推論を繰り返すベイズ更新は、〈環境の現実には原理的に知り得ないため、定まった値には到達し得ず、いつまでも確率計算を繰り返す他ない〉という前提に立っている。このようにベイズ統計は確率的性格が強いため、確率に依拠する熱力学や情報理論に関わるエントロピーや自由エネルギーが、ベイズ統計にとって重要な概念となっているのである（渡辺 二〇一二：八）。

実際、脳の作動もベイズ推論の前提に立っているように思える。例えば視覚は眼球の網膜上の二次元画像を感覚信号として用いるが、その精度はデジカメより遥かに劣る（乾・阪口 二〇二一：一四）。それにもかかわらず、網膜上の二次元画像から外界の三次元イメージを作り上げている。つまり、少ない感覚信号から、大幅な推論の力を借りて、豊かな外界像を脳内に構築していることになる。

ベイズ更新から窺えるように、ベイズ推論の特徴の一つは「逐次的合理性」である（小島 二〇一五：一三八、一四四）。ベイズ更新では、感覚信号の観測が行われる度に、観測結果に応じて尤度が更新されるので、計算が再帰的にやり直され、新しい推論結果が得られ続ける。合理性は一挙に獲得されるのではなく、逐次的に形成されていく。この場合の合理性は、時間を越えた永遠不変の合理性ではなく、あくまで時間の流れの中での合理性である。全体が一挙に与えられるのではなく、部分が一つずつ積み重なって、結果的に全体が事後的に形成される。

トーマス・ソームズは脳内のベイズ推論を、ダムに問題がないかを検査するスタッフの寓話で説明する。ダム施設は巨大で複雑なので、個々の検査員はその全貌を知り得ないし、ダム全体の目的も知り得ない。だが、全体を知らない検査員たちはそれぞれに割り当てられた局所的な観測・調整作業を次々と実施し続けるだけで、全体の目的が適切に設定され遂行されていく（ソームズ 二〇二一：二三三、ホーヴィ他 二〇二一：九六も参照）。これも一種の逐次的合理性だろう。

さらに言えば、諸部分が逐次的に積み重なると言っても、ブロック玩具遊びのように、縦方向や横方向へ空間的に組み上がり広がっていくのではなく、時間の中で再帰的に積み重なっていく。ベイズ更新を繰り返した後に現れるのは、複数の事後確率の群ではなく、たった一つの事後確率である。だが、その数値には、ベイズ更新の度に、新たな感覚信号が繰り返込まれ続けている。それら時間的に再帰的な積み重なりが、最後の一つの事後確率として結実している。

私見では、逐次的合理性の有り様は、美術におけるデッサンの作業に似ている。最初は大雑把な当たりを付けて取り敢えず線を引き始めるのだが、その後次第に、先に引いた線を手掛かりにして逐次的に、より精確な線へと収斂していく。デッサンの作業を繰り返すうちに、最初に引いた線は後から引いた他の線に紛れてしまうかも知れないし、消されてしまうかも知れない。だが、たとえ好い加減であっても最初の線（事前信念）が引かれなければ、後の線は引けないだろう。精密な下絵を予め完成させてから本画に取りかかるやり方が古典統計に似ているとするならば、ベイズ更新はデッサンの営みに似ている。

以上のような推論が脳で作動しているのではないかと、自由エネルギー原理は想定している。ただし、実際の脳は厳密なベイズ推論をしていないと考えられる。なぜなら、厳密な計算のために必要な数値が現実には入手出来ないからである。脳が推論しなければならない外界という環境は、〈壺の中に白と赤の玉がいくつか入っていて、その中からランダムに一個を取り出す〉というようなモデルとは大きく異なっている。この種の壺のモデルでは、赤い玉を取り出す確率を一義的に、一つの決まった数値として計算出来る。だが、脳が直面する外界は通常そのようなモデルでは出来上がっておらず、敢えて言えば、中にどんな物が何個入っているか見当の付かない壺のようなものである。つまり、確率を割り算で計算する際の実母が分からない。

そこで、厳密なベイズ推論の代わりに近似計算せざるを得ない。だが、近似計算によって導き出される事後確率（近似事後確率）はあくまで近似なので、真の事後確率とは距離がある。その

距離を測るための方法として、自由エネルギー原理では、カルバック・ライブラー・ダイバージェンス (KLD) を用いる。KLD は「観測から外界を「最適に」推定する方法」において重要な役割を果たす (田中 二〇一九：九三、渡辺 二〇一二：九も参照)。KLD が小さいほど、近似事後確率は真の事後確率に近いのである。

だが、KLD によって真の事後確率との距離を計算しようとする、真の事後確率が既に分かっている必要があるが、既述のように、真の事後確率は原理的に知り得ない。

ところが、KLD 計算式にベイズの定理を用いると、 $p(x, y)$  と  $p(y)$  が導入される代わりに、真の事後確率  $p(x|y)$  を消去出来る。そして式を整理した後で最後に現れるのが自由エネルギー項とサプライズ項である (乾・阪口 二〇二一：三〇)。運動と関わりのあるサプライズ項は、感覚信号とは関わるが、近似事後確率には依存しない (乾・阪口 二〇二一：六二)。よって、当面サプライズ項を棚上げにすれば、自由エネルギー項が最小の時に KLD は最小となり、それは、近似事後確率が (脳が原理的に知り得ない) 真の事後確率に最も近づく状態である。このようにして、近似的ながらベイズ推論は可能となる。

つまり、厳密なベイズ推論が不可能という現実条件下で近似的にベイズ推論するためには、自由エネルギー最小化を判断基準にすれば良い。脳にとって原理的に知り得ない外界という環境をそれでもなお脳が把握するための鍵は、自由エネルギーの最小化なのである。そして自由エネルギーは、観測者にとって把握しがたいエントロピーよりも扱いやすい量だと言う (鈴木 二〇一四：二〇二～二〇三)。

アニル・セスによれば、「通常、ある量を最小化するためには、システムはそれを測定できないなければならない。しかし、ここで問題となるのが、感覚的なエントロピーは直接的に検出したり測定したりすることができないということである」。だが「私たちの目的では、自由エネルギーは感覚的なエントロピーに近い量と考えることができる」。そして自由エネルギーはエントロピーと違い「生物によって測定可能であること、それゆえ、生物が最小化できるものである」。ゆえに「生物は「自由エネルギー」という測定可能な量を能動的に最小化することによって、自らの生存を保証する低エントロピー状態を維持していることになる」(セス 二〇二二：二二二～二二三)。このように自由エネルギー最小化を基本として、脳の様々な作動を統一的に説明していこうとするのが自由エネルギー原理である。

このように見てくると、脳におけるベイズ推論は、厳密な計算としては成り立たず、遠回りをして辛うじて近似的に成り立っているに過ぎないようにも見える。つまり、確固たる現実を把握するには、ベイズ推論が弱々しい道具に見えてしまうかも知れない。だがそれは、既述の「常道」から見た風景である。逆方向の「王道」から見れば、自由エネルギーの最小化こそ真の基本原理であり、私たちがベイズ推論と呼んでいるメカニズムやその結果としての「定型発達者」の現実感とは逆に、自由エネルギー原理から近似的に作り出された構築物と言えるのではないだろう

か。現実を精確に把握するという分かりやすい目標よりも、自由エネルギーを最小化することの方が根本的なのである。それはすなわち、現実を精確に把握するために自由エネルギーを最小化するのではなく、自由エネルギー最小化のために現実を把握するということになるのかも知れない。

自由エネルギーの最小化は熱力学や情報理論の基礎である。化学反応の方向を決める要因も自由エネルギーである。それを思えば、このような統一的原理の主張は、脳だけを特別に考えがちな私たちの常識を揺さぶる大胆で魅力的なアイデアである。「王道」の行き着く先は、社会性をも自由エネルギーの最小化で説明することなのだろうか。だがこれを反証可能な形で検証するのは難しい（セス 二〇二二：二二五）。今後、この原理による計算モデルやシミュレーションが現実の観察とどこまで一致するか繰り返し調べる必要があり、その研究過程自体がベイズ更新的営みとなるだろう。

### 三 自閉症と一覽表的合理性

では、自由エネルギー原理から近似的に導き出されたベイズ推論がうまく作動しなかった場合はどうなるのか。それが、統合失調症や自閉症なのではという主張がある（田口 二〇一九：二二、ホーヴィ 二〇二一：二五一）。それでは、ベイズ更新にはどのようなタイプの不具合が考えられるのか。事後確率  $p(x|y)$  を導き出すベイズの定理が事前確率  $p(x)$  と尤度  $p(y|x)$  から出来上がっているのだから、不具合の一つのタイプは、事前確率すなわち事前信念が計算上、過大あるいは過小に扱われる場合が考えられる。そして、もう一つのタイプは、尤度を構成する要素である感覚信号が計算上、過大あるいは過小に扱われる場合である。

自由エネルギー原理による統合失調症や自閉症のメカニズムの説明は発展途上だが、現時点では、統合失調症では、事前確率つまり信念がベイズ推論に過大な影響を与えているのではと推測されている（乾・門脇 二〇二三：一三五）。外界を無視したような幻覚や妄想は、信念が外界からの感覚信号によって適切に調整されていないため、信念の暴走状態に陥っているのではないかという考え方である。

事前信念は感覚信号の観測以前から脳内に存在するが、そもそも脳はいかにして最初の信念を形成するのか。ベイズ推論では、事前確率の設定に関して、「理由不十分の原理」が想定される場合がある（小島 二〇一五：四四）。つまり、ある事象が起きるかどうかについて、その是非を判断するための根拠となる観測が得られない場合、起きる確率を五〇%、起きない確率を五〇%として取り敢えず設定する。

この理由不十分の原理を幻覚や妄想について応用してみると、極端に言えば、途方もない妄想であっても、それが現実であるか否かを判断するために役立つ観測事実が得られていない初期段階では、妄想を裏付ける出来事が実際に起きている確率は五〇%だと取り敢えず見なすことにな

るだろう。もちろんこれは非常に高い確率だが、観測が逐次的に繰り返されるに連れて急速に低くなっていくはずである。これが定型発達者にとっては馴染みの状況である。

だが、もしベイズ更新が適切に繰り返されずに、「理由不十分の原理」から生まれた妄想が現実である確率が順調に低くならなかったらどうか。重要なのは、これが単なる計算間違いの不具合ではないという点である。尤度だけでなく事前確率をも考慮に入れることは、その事前確率（主観確率とも言う）がたとえ非現実的であっても、ベイズ推論において間違いではない。ただ、更新という再帰性が継続することが重要なのであり、順調な更新無しでは、非現実的な事前確率の数値を下げて妄想を抑えるのは容易ではない。このような不具合のあり方は、古典統計では想定しにくい、ベイズ統計では想定しやすいのではないだろうか。

アニル・セスによれば「[「正常な」知覚も「異常な」幻覚も、どちらも感覚入力の原因について内的に生成された予測であり、脳の中核的なメカニズムを共有している。しかし、「正常な」知覚の場合は、知覚したものが世界における〔その知覚の〕原因〔と想定されているもの〕と結びついており、それによって制御されているのに対し、幻覚の場合は、知覚したものがある程度、その原因に対する支配を失っているという違いがある」（セス 二〇二二：九九、竹中 二〇一二も参照）。

それに比べて自閉症の場合、感覚信号が過大に扱われているのではないかと推測されている（乾・門脇 二〇二三：一四〇）。自閉症には、そのネーミングとは裏腹に、外界からの影響を過大に受け取ってしまう傾向がある。例えば感覚過敏がそうだが、それに限らず、他者の発言を字義通りに受け取ってしまうこともその一面かも知れない。しかしだからと言って、自閉症者の感覚器官の機能が生理的に過敏になっているわけではないようなので、感覚信号を脳内でどのように処理しているかが問題なのだろう。

感覚信号は、尤度を構成する一部である。もし自閉症者のベイズ推論において、事前確率ではなく尤度が過大に影響しているならば、感覚信号の過大評価ということになる。そして新たなベイズ更新の度、感覚信号は様々でありうる。この状態でベイズ更新が繰り返されれば、変動が比較的少ない事前確率よりも、観測ごとに変動が大きい感覚信号の影響が過大にベイズ推論に影響することになる。結果的にこれが感覚過敏等の困難を引き起こすのではないだろうか。

感覚過敏の困難については、〈経験する回数を増やせば、次第に慣れていくのではないかと〉疑義を抱く人がいるかも知れない。しかし、もしベイズ更新のメカニズムがうまく行っていないのならば、回数を重ねることは必ずしも困難の緩和には繋がらない。経験値を積むことで状況が改善するためには、ベイズ更新が適切に働いて逐次的合理性が適切に作動しているという前提がまず必要なのである。

部分と全体の関係で言えば、感覚信号が適切に評価されてベイズ更新が進行すれば、部分の多少の変化は全体に決定的影響は与えないだろう。だが、感覚信号が過大に評価されるなら、部分

の些細な変化が全体に大きな影響を与えてしまう。例えば、人の顔を認識する際、ほくろの有無ぐらいは大きな影響を与えないはずだが、そうではない自閉症者が存在する。定型発達者の顔認識は、〈部分における変化が多少あったとしても全体に対して大きな影響は与えないはず〉という前提があって初めて成り立つのである。

自閉症者が「自閉」的に見えるのは、当事者自らが閉じようとしているというよりも、外界からの感覚信号の影響が甚大になってしまうため、その大嵐から身を守るために閉ざさざるを得ないからではないだろうか。規則や時刻を過剰に遵守しようとするのも、外界における想定外の変動に影響されないための必死の防御策であると思われる。

一見すると自閉症者の頑固さは、事前信念が強すぎるためと見えるかも知れない。だがむしろ話は逆ではないだろうか。自閉症のある児童生徒が教室で快適に過ごせるよう、教室内の物や情報を整理整頓し少なくするという「構造化」手法はこの点で示唆的である。もし事前信念が強すぎるのが困難の原因ならば、外界の整理整頓がその困難の緩和に役立つはずはない。尤度や感覚信号の方に問題があるからこそ、本人の心構えを変えようとする無理な努力を期待するよりも、環境を整備する方が合理的なのである。

自閉症療育において、一日のスケジュールを明確に示した一覧表が活用される。山頂から下界の風景を一望の下に眺めるように、一日の時間全体を眺められるスケジュール表は、時間の流れの中で生じる逐次的合理性がうまく作動していない自閉症者に心の安定をもたらすようである。もし逐次的合理性がうまく作動していれば、時間の流れの中にいるのはそれほど不安ではないだろうが、うまく作動していない場合、不安に駆られ、時間を越えた全体秩序を一挙に手に入れたいと思ってしまうだろう。時間の流れを空間の広がりへと擬似的に変換したスケジュール一覧表は、ベイズ統計的と言うよりは古典統計的なのである。

自閉症研究において有名な「サリーとアン課題」もベイズ推論の不具合との関連で説明可能な面がある。サリーがビー玉を籠の中に入れた後その部屋を離れている隙に、いたずらなアンがビー玉を箱の中へ移してしまう。では、サリーが部屋に戻ってきて、ビー玉を取り出そうとする際、どこを探すだろうか。この答を被験者に問うのが、「サリーとアン課題」である。先行研究によれば、被験者が自閉症である場合、「箱の中」と答える傾向があるとされている。

「サリーとアン課題」を下支えする「心の理論」はいかにも限定的で実験室的であり、様々な批判を受けているのは事実である。だが、批判論集の監訳者・藤野博も、他ならぬ自閉症の文脈で「心の理論」を一定程度評価しているのは示唆的である（Leudar 他編 二〇二三：三六三）。

自閉症者にとってこの課題が難しいのは、〈アンや被験者が既に知っていること（ビー玉は箱の中にあること）〉をサリーは知らないという点である。空間的で一覧表的な把握をする人にとって、ビー玉が箱の中にあるのは自明なので、サリーだけが〈ビー玉は籠の中にある〉と思っているとは思えない。実験場面の上空から写真を撮れば、サリーもアンも、箱の中にビー玉があると

いう状況と共存すると被験者には見えるだろう。だが、サリーの中では常に逐次的合理性実現のためベイズ更新が再帰的に作動し続けている。この点に被験者が気付けば、ビー玉が箱へ移された場面を感覚信号として取得していないサリーだけが〈ビー玉が今や箱の中にあること〉を知らないと推論出来るだろう。

だが、もし被験者が、予想し得ない感覚信号の襲来に対する防御として、一覽表的合理性に従っているとしたらどうか。逐次的合理性は、次々と到来する感覚信号をその度に逐次的に繰り込んで推論を更新していく点で柔軟だが、その代わり、次々と訪れる感覚信号によるサプライズに耐えねばならない。だが自閉症者の場合、ベイズ推論において感覚信号が過大に作用するのであれば、そのサプライズは耐えがたく、逐次的合理性とは別の合理性、つまり一覽表的な合理性を求めようになるのかも知れない。

「サリーとアン課題」の場面を時間的ではなく空間的に眺めてしまうと、サリーの内面で起こる逐次的で再帰的な変化は把握しにくい。被験者が場面を空間的な一覽表として眺めてしまうと、〈アンや被験者自身はビー玉を箱へ移した場面を視覚的な感覚信号として受け取ったが、サリーだけはその感覚信号を受け取らずに独自のベイズ更新をしている〉という全体状況を把握しにくいのである。

このようにベイズ統計は、逐次的で再帰的な状況の把握に適している。さらにこの特徴は、「世界の分岐」という考えとも関わる。小島寛之はベイズ統計の特徴として「世界の分岐」を挙げる(小島 二〇一五：一九)。ベイズ更新の度に新たな感覚信号を考慮に入れることで、時間の流れの中で、それ以前とは別の世界が次々と出現していく。サリーがいない時にアンがビー玉を移動させたために、アン(と被験者)の世界とサリーの世界とは分岐したのである。

このような「世界の分岐」は、空間における分岐ではない。分岐した複数の世界は空間的に区分けされるわけではない。アン(と被験者)の世界とサリーの世界とは空間上で並列してはいない。むしろ時間軸が分岐したのであって、ゆえに世界の分岐は空間的一覽表の眼差しでは捉えにくい。

逆に言えば、このような眼差しを持つ人にとっては、空間的な区分の方が理解しやすいということになる。実際、自閉症療育において、空間の視覚的区分を自閉症者の心の安定に役立てている。例えば、同じ場所で時間帯ごとに異なる作業をするより、作業ごとに異なる空間へ移動する方が不安を引き起こしにくいと考えられている。これもまた、時間の空間化の一例だろう。それによって、本来は時間的性質が強い「世界の分岐」は、擬似的にだが空間的に把握出来るようになるのである<sup>(1)</sup>。

#### 四 自閉症者の道徳教育への示唆

日常生活において、逐次的合理性が大きな意味を持つ場面は多い。「サリーとアン課題」はあ

くまで実験室的に設定された特殊状況であるが、以下では、実際に自閉症者が日常的に経験する場面の一つとして、学校教育の場における合理性について考えてみよう。

学校教育の場における逐次的合理性の重要性は、教科によって差がある。例えば、算数や数学では逐次的合理性の発揮の機会は比較的少ない。数学的証明の過程は一見すると時系列的な展開に見えるが、その内実は空間的な一覧表だと言えそうである。それに比べ国語においては、逐次的合理性の比重は大きい。小説の登場人物の気持ちを問う問題に答えるには、その人物の事前信念を出発点とした上で、次々に起こる出来事をきっかけにして新規に受け取り続ける感覚信号が登場人物の事後信念をいかに再帰的に変容させていくか、その過程を把握しなければならない。自閉症のある児童生徒が、教科としての国語とりわけフィクション読解を苦手としがちだという知見がある。

そして、国語以上に苦手と推測されるのが、道德教育である（永田監修 二〇二一）。もしもそれが既存の社会的ルールの遵守を習得するだけの内容ならば、自閉症のある児童生徒は理科や社会科と同程度に対応可能だろう。しかし、「特別の教科 道德」はそのような内容ではない。事前に決まった答を探し出すのではなく、再帰的営みによる逐次的合理性の発揮が求められている。

「二通の手紙」（原作は白木みどり）という授業教材を取り上げてみよう（杉中 二〇二三：四九）。動物園で働く元さんの下へ、入園時刻を過ぎてから幼い姉弟が訪れ、弟の誕生日祝のため入園を懇願する。二人の事情を知った元さんは取返して規則を破って入園させたが、その後、なかなか戻ってこないため騒ぎとなる。二人は無事見つかるが、元さんは叱責を受ける。後日、姉弟の母親から感謝の手紙が届くが、その翌日、停職処分通知の手紙が元さんの下へ届く。元さんはすっかりした様子で退職を選ぶ。

上記のように二通の手紙は、日付を違えて送られてくる。この時間差が重要である。そもそも、全ての出来事は一挙に、一覧表のように生じたのではなく、逐次的に起こった。姉弟が元さんの下へやってきた時には、後に騒ぎが起きるとは想定されていなかったし、母親の手紙は、それ以前に姉弟が無事に見つかったという事実を前提としている。このような一連の時間の流れの中で、姉弟が元さんの下へやってきた時点で彼が下した判断を、その時点の状況に基づいて評価するか、それとも、それ以降に起こった出来事も考慮に入れるかによって、評価自体が異なってくる。そこには逐次的合理性の問題が横たわっている。

文脈は異なるものの、自閉症研究における「サリーとアン課題」でも、登場人物二人が持つ知識が逐次的に変化していく流れを適切に把握し、二人が逐次的合理性を駆使して、各時点で自分が持っている知識だけに基づいて判断していることを被験者自らも逐次的合理性に従って理解して初めて、正しく解答出来るのである。

姉弟が動物園を訪れた時に元さんは、その後に起こる一連の出来事をまだ知らない。元さんの判断と行動は、そのような背景においてなされた。しかし読者は、何度も物語を読み返せる立場

にあるので、元さんが知っていることと知らないこと全体を一覧表的に眺望可能である。元さんの立場は「サリーとアン課題」におけるサリーに似ており、読者の立場はこの課題の被験者に似ている。

もし自閉症のある児童生徒ならば、「二通の手紙」の中で起きる全ての出来事を一覧表上に横並びさせて、〈規則違反は規則違反だ〉として明快に評価するかも知れない。しかしながら、道徳の授業でこの教材を用いる趣旨は、単に遵法精神を再確認するためではなく、単純明快な可否判断では収まりが付かない場面が世の中にあるのを気付かせることだと思われる。そんな割り切れなさが、道徳教育の難しさであると同時に、この教科の特徴や意義でもある。

だが、自閉症のある児童生徒の場合、そのような授業の趣旨を容易に理解出来ない可能性がある。脳の機能障害のせいで、逐次的合理性を実現するベイズ更新が不具合に陥っている児童生徒が教室にいるのなら、道徳教育のためには何らかの特別な工夫が必要になってくるだろう。少なくとも、教材理解の困難をその生徒の道徳意識の低さのせいにはならないはずである。

古典統計とベイズ統計の間には確執があり今日に至っているが、それは技法上の違いだけでなく、曖昧な状況の中で何をどう信頼するかをめぐる世界観の違いでもある。道徳教育をめぐる自閉症者と定型発達者の間の擦れ違いにも、それと相似形の課題が横たわっているように思われる。

同時代の偏狭な道徳意識の犠牲となった数学者アラン・チューリングは自閉症のだったのではと今では推測されている。彼は二次大戦中、ナチスの暗号生成機「エニグマ」(謎)の解明に大きな貢献をなした(マグレイン 二〇一八:一六一)。戦後長く機密とされてきたこの事実は近年広く知られるようになったが、彼は、立ちほだかる謎を解読する際にベイズ推論を用いたと言われている(豊田秀樹編著 二〇一五:二〇)。もし自閉症者が脳内におけるベイズ推論の困難ゆえに社会性の謎を前にして立ちすくんでいるのだとすれば、このエピソードは現代においても示唆的である。

#### 注

- (1) この注自体が本論における「世界の分岐」である。本文においては、児童生徒が学ぶべき道徳の存在を前提として論じたが、もちろん、このような定型発達的前提自体を括弧に入れる立場もありうる。内海健は、ドゥルーズが用いる対比(真の反復・経験・差異／交換・習慣・誤差)を援用して、自閉症者のあり方を論じている。この一連の対比の後者は「トップダウンの認知」(内海 二〇二〇:二〇〇)と結び付いている。ドゥルーズは、後者に位置付けられる「表象」を、前者に位置付けられる「生成」に対する抑圧装置として捉えた。つまり、後者の「表象」の下、前者の「差異」は、単なる「変異や誤差に過ぎ」ないと見なされ、結果的に(「真の反復」とは違って)固定的な「モデル」へと収束される。教育されるべき道徳もまたモデルの一種だろう。内海によれば、同一性保持のために反復する自閉症者の「強度的(intensive)」なあり方は前者に属する。この際の強度的とは、「世界の局外に立つ(ex-ist)」ことが出来ず「外部尺度が持ち込まれない状態」である(内海 二〇二〇:一九五)。自閉症は「徹底的に内在的な生のいとなみ」であり(内海 二〇二〇:一九六)、内部／外部という対が未だ成立しておらず、「世界の亀裂を与えるもの」としての「他者」がない状態なのである(内海 二〇二〇:一九五)。

「徹底的に内在的」というのは、内部に留まるという意ではなく、「構造的に外部というものをもたない」状態である（内海 二〇二〇：二一〇）。内部が先に生まれたのでもなければ、外部が先に存在したのでもない。「前もってある「地」の上に「図」が浮かび上がるのではなく、「地」も「図」と同時に設定されるということである」（内海 二〇二〇：一九八）。強度的なあり方においては未だ「他者は外部ではない」（内海 二〇二〇：二一〇）。

定型発達者のように、もしも内部／外部という対の「局外に立つ」ことが出来れば、反復の営みなどに頼らなくても同一性を保持出来るだろう。なぜなら、「局外に立つ」ことが出来れば、反復は高々オリジナル（ある意味で「モデル」）のコピーに過ぎず、オリジナルほどの値打ちを持ちえないからである。しかし自閉症者の強度的なあり方においては、反復でしか同一性を保持出来ないとも言えるし、見方を変えれば、「一回的偶然をそのつど反復する永遠」のような経験」（内海 二〇二〇：二一三）として反復を高く評価することも可能だろう。

このような自閉症者の反復が「病理」的で「適応に不利」なのを認めた上で、内海はそこに「TD（定型発達：引用者注）を支配している表象体制を揺さぶる固有性、創造性」を見出そうとしている（内海 二〇二〇：二一一）。

以上の内海による理解は、本文で紹介した自由エネルギー原理と相性が良いように思う。自由エネルギー原理によれば、定型発達ではベイズ推論によって、トップダウンの事前確率とボトムアップの感覚入力（ドゥルーズの言う「差異」）が合わせて計算された結果、穏当な事後確率となる。内海の表現を使えば、そこにおいて「差異」は「所詮はちょっとした変異や誤差」へと馴化されてしまい、差異を穏当に収束させる「モデル」として固定化するのである。

それに対し自閉症者では、トップダウンの事前確率の影響力は相対的に弱く、ボトムアップの感覚入力の影響力が強い。それゆえ自閉症者の世界は強度的なものである。ドゥルーズの強度的主張は間接的ながら、定型発達への批判の契機としての自閉症の可能性を示唆しているのではないだろうか。ところで、ドゥルーズが用いる強度（内包量）／外延量という対比は、自由エネルギー概念とも密接な関係にある（内包量／外延量、intensive / extensive については竹中 二〇一〇も参照）。

熱の出入りを伴う化学反応が起きつつある容器があるとしよう。熱力学第二法則により、化学反応はエントロピーが増大する方向へ進む。だが第二法則が示しているのは、容器内部だけのエントロピーの増大ではなく、内部と外部のエントロピーの合計が増大するということである。したがって化学反応がどの方向へ進むかを知るには、容器の内部と外部のエントロピーの合計を計算する必要がある（鈴木 二〇一四：一九九）。とは言うものの、外部（容器以外の宇宙全体）のエントロピーは到底計算出来そうもない。本文で言及した〈エントロピーは把握しがたい〉という点はこれと関わる。ところが、「容器の外で起こっているエントロピー変化」は、容器から外部へ流れ出た熱エネルギーを、容器内部の絶対温度（以下、温度）で割り算すれば計算出来るのである（鈴木 二〇一四：二〇一）。つまり、容器内部の温度が分かれば「容器の外の世界は一切気にしなくていいことになる」（鈴木 二〇一四：二〇二）。このように、容器内部のエネルギーから、内部のエントロピーと温度をかけたものを引き算したのが自由エネルギーである。

つまり、外部を観測せずとも内部の情報だけによる計算で、内部と外部の全体を把握出来る。「神ならぬ身のわれわれは宇宙を一望することが叶わぬから全体のエントロピーを把握することはできない。それでも自由エネルギーというデバイスを駆使すれば、目の前の反応を理解し予言することができる」（鈴木 二〇一四：二〇三）。自由エネルギーがエントロピーより扱いやすいとされる所以である。

それは、自由エネルギー計算式に温度という強度（intensity）が含まれるからだろう。エネルギーやエントロピーが示量（extensive）変数であり、内部と外部の区別に従うのに対し、温度という示強（intensive）変数は内部と外部を区別しない。温度は強度、内海の表現を借りれば「外部尺度が持ち込まれない状態」なのである。エネルギーやエントロピーのような示量変数だけの計算では、内部は知り得るが外部を知り得ない。だからこそ、示強変数と示量変数が掛け算された結果（共役関係）としての自由エネルギーが重要である。

エントロピー等の新しい概念に目が行きがちだが実は、温度という旧来の概念こそ驚きなのである（鈴木 二〇一四：二〇一～二〇五）。エントロピー原理ではなく自由エネルギー原理である所以はそこにある。

内部エネルギーやエントロピーが持っていない、自由エネルギーのこの性質が、内海の論点（内部と外部が同時成立するプロセス、言い換えれば自己と他者の同時成立）と相通じているのは偶然だろうか。外部を知ることが原理的に出来ないにもかかわらず、外部のことを知らねばならないというのは、脳が強いられている条件と同じである。自由エネルギー原理の提唱はまだ始まったばかりである。だが、もし自由エネルギー原理が正しいとすれば、それは、他者による社会的自己の成立という社会学的課題を、温度という示強変数が重要な役割を果たす熱力学や統計力学を駆使する計算論的神経科学によって説明する試みの端緒と言えないだろうか。

#### 参考文献

- 乾敏郎・阪口豊、二〇二一、『自由エネルギー原理入門—知覚・行動・コミュニケーションの計算理論』岩波書店。
- 乾敏郎・門脇加江子、二〇二三、『脳科学はウェルビーイングをどう語るか』ミネルヴァ書房。
- 内海健、二〇二〇、「反復と強度」、内海健・清水光恵・鈴木國文編『発達障害の精神病理Ⅱ』星和書店。
- 小島寛之、二〇一五、『完全独習 ベイズ統計学入門』ダイヤモンド社。
- 杉中康平、二〇二三、「二通の手紙」を通して考える道徳科教材論（授業論）～「生き方」についての考えを深める道徳科授業を目指して』『道徳教育論叢』第一号、日本道徳教育方法学会。
- 鈴木炎、二〇一四、『エントロピーをめぐる冒険—初心者のための統計熱力学』講談社。
- セス、ア Nil 著、岸本寛史訳、二〇二二、『なぜ私は私であるのか—神経科学が解き明かした意識の謎』青土社。
- ソームズ、トーマス著、岸本寛史・佐渡忠洋訳、二〇二一、『意識はどこから生まれてくるのか』青土社。
- 田口茂、二〇一九、「ベイズ理論から帰結する現実観の変容—浅井論文へのコメント」『心理学評論』六二（一）。
- 竹中均、二〇一二、『精神分析と自閉症—フロイトからヴィトゲンシュタインへ』講談社。
- 田中宏和、二〇一九、『計算論的神経科学—脳の運動制御・感覚処理機構の理論的理解へ』森北出版。
- 豊田秀樹編著、二〇一五、『基礎からのベイズ統計学—ハミルトニアンモンテカルロ法による実践的入門』朝倉書店。
- 豊田秀樹、二〇一六、『はじめての統計データ分析—ベイズ的〈ポスト p 値時代〉の統計学』朝倉書店。
- 永田繁雄監修、齋藤大地・水内豊和編著、松尾直博・細川かおり著、二〇二一、『新時代を生きる力を育む 知的・発達障害のある子の道徳教育実践』ジアース教育新社。
- パー、トーマス、ジョバンニ・ベッツォーロ、カール・フリストン著、乾敏郎訳、二〇二二、『能動的推論—心、脳、行動の自由エネルギー原理』ミネルヴァ書房。
- ホーヴィ、ヤコブ著、佐藤亮司監訳、太田陽・次田瞬・林禪之・三品由紀子訳、二〇二一、『予測する心』勁草書房。
- マグレイン、シャロン・バーチュ著、富永星訳、二〇一八、『異端の統計学 ベイズ』草思社。
- Leudar, Ivan & Alan Costall 編、熊谷晋一郎・藤野博監訳、二〇二三、『「心の理論」は必要か—心のありかを探る12の視点』金子書房。
- 渡辺澄夫、二〇一二、『ベイズ統計の理論と方法』コロナ社。

付記 本稿は科研費補助金・基盤研究（C）「学校における発達障害の知と経験に関する研究—道徳教育を中心に」（課題番号：24K05295、研究代表者：竹中均）の成果の一部である。また、早稲田大学特定課題研究助成 2025R-011を受けている。