

# 拡散過程モデルによる潜在的連合テスト（IAT）データ分析の実際

土 居 淳 子  
川 西 千 弘

## I はじめに

土居・川西（2012）で Ratcliff の拡散過程モデルとパラメータ推定法を紹介し、潜在的連合テスト（Implicit Association Test、IAT）で得られる正答および誤答の反応時間データの分析に適用することの可能性について述べた。その後、土居・川西（2014）において、母親 258 名の子どもや養育に対する被害（加害）意識を IAT によって測定したデータを分析し、土居・川西（2015）では、Ratcliff の拡散過程モデルに基づきパラメータ推定するための分析パッケージ開発について報告した。

本稿では、IAT と拡散過程モデルの要点を簡単にまとめた上で、IAT データを分析するための具体的なノウハウと、実データの分析結果を紹介する。

## II 反応時間の拡散過程モデルと IAT

### 1. Ratcliff の拡散過程モデル

まず、本研究で用いる Ratcliff（1978）の拡散過程モデルについて簡単に述べる。このモデルでは、二肢選択分類課題において、刺激が提示されてからどちらかを選択するまでの意思決定プロセスを一次元の情報（確証）蓄積プロセスであると仮定する。刺激が提示されることによって、2つの選択肢のどちらを選択すべきかに関する情報量が時間経過とともに増減し、どちらかの選択肢を選択するに足る情報蓄積量に達したときに、それを選択するという意思決定が成されると考える。

この情報蓄積プロセスに、さまざまな発生源からもたらされる「ゆらぎ」が含まれることを仮定すると、提示される刺激が同じでも、試行ごとに選択までの時間（反応時間）は一定でなくなり、大きなゆらぎにさらされた場合には、通常とは逆の選択が生じる場合もある。Ratcliff の拡散過程モデルでは、刺激提示後の情報蓄積プロセスを、一定のドリフトと不規則なブラ

ウン運動の組み合わせである Wiener 過程であると仮定し、刺激が提示されてから選択を行うまでの時間を、ある初期値から出発した Wiener 過程が上下どちらかの閾値に達するまでの初通過時間でモデル化する。同一の刺激呈示によって同一のドリフトにさらされたとしても、ランダムなゆらぎによって、閾値に至るまでの情報蓄積の時間経過は試行毎に異なるので、試行毎に異なる反応時間が得られる。

図 1 は、その概要を図式化したものである（横軸：時間経過、縦軸：情報蓄積量）。Wiener 過程のドリフト  $v$  は単位時間当たりの情報収集率に相当し、無限小分散  $s^2$  はブラウン運動の強さ、いいかえると、ゆらぎの強度を表す<sup>1)</sup>。その他の主要パラメータは閾値間の距離  $a$ 、初期値  $z$ 、符号化や運動に要する非決定時間  $t_0$  の 4 つで、 $s^2$  を基準とした相対的な値を表す。ドリフト  $v$  は、情報処理スピードを表すパラメータで、難しい課題の場合ほどその値は小さくなる。また、被験者間の比較を行う場合には、各人の情報処理スピードの指標となる。閾値間の距離  $a$  は、その値が大きいほど決定までの時間は長くなり（平均として）、かつ、ランダムなゆらぎによる間違った選択が成されにくくなる。したがって、判断の慎重さを表すパラメータである。初期値  $z$  は決定バイアス（decision bias,

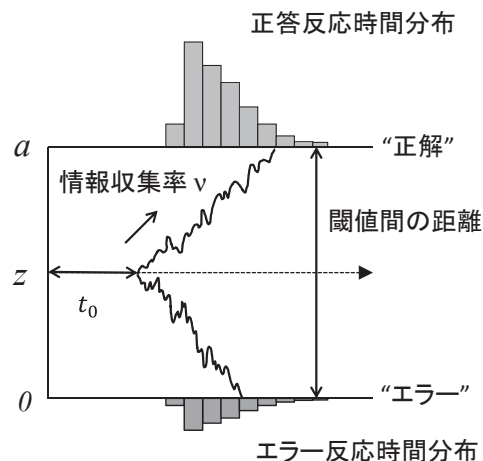


図 1 Ratcliff の拡散過程モデル

response bias) と呼ばれるもので、初期値  $z$  がいずれかの閾値に近いほど、その閾値に対応する選択が成されやすくなる。決定バイアスがない場合、 $z=a/2$  である。非決定時間  $t_0$  は、符号化や運動に要する等、決定プロセスに付随するその他の時間をすべて足し合わせたパラメータである。

さらに、Ratcliff&Rouder (1998) で、初期値  $z$ 、非決定時間  $t_0$  および情報収集率  $v$  を確率変数とするモデルが提案され、それぞれのバラツキの大きさを表す3つのパラメータ ( $s_z, s_t, s_v$ ) が追加された。初期値は  $z$  を中心とする幅  $s_z$  の一様分布、非決定時間は  $t_0$  を中心とする幅  $s_t$  の一様分布、情報収集率は平均  $v$ 、標準偏差  $s_v$  の正規分布に従うと仮定する。つまり、初期値・非決定時間・情報収集率の値が試行毎に異なる値をとり得ることをモデル化したのである。この拡張によって、エラー回答の反応時間のばらつき等、反応時間の分布をより正確に再現できるようになった。このように、Ratcliff の拡散過程モデルには7つのパラメータが存在する (表1)。

表1 拡散過程モデルのパラメータとその解釈

記号	パラメータ	心理学的解釈
$a$	閾値間の距離	慎重さ
$v$	情報収集率 (平均値)	情報処理の速さ
$z$	初期値 (平均値)	先験的なバイアス
$t_0$	非決定時間 (平均値)	運動や符号化に要する時間
$s_v$	情報収集率の標準偏差	
$s_z$	初期値 $z$ の変動幅	
$s_t$	非決定時間の変動幅	

## 2. 潜在的連合テスト (IAT) と IAT スコア

まず、IAT の具体的な内容を簡単に紹介しよう。

ある概念のペア (例えば、「花」と「虫」) とそれらの連合の対象となるもう1組の概念ペア (例えば、「快」と「不快」) を想定し、それぞれの概念に属すると考えられる単語や絵、写真などの表象を刺激として用意する。実験参加者は、コンピュータ画面中央に1つずつ表示される刺激を、右ないし左の概念へと出来るだけ早く分類することが要求される (図2)。画面上部の左右には概念名が表示されており、参加者は右または左に対応するキーを押すことにより、それぞれの刺激がどちらの概念に属するかを分類していく。

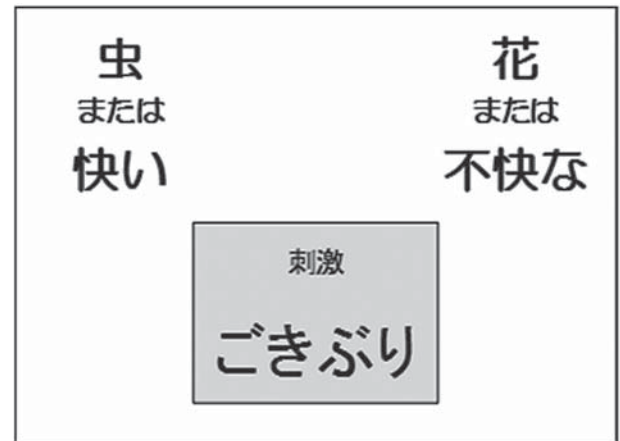


図2 IAT 画面の模式図

このような弁別課題において、「花」と「虫」に属する単語を弁別する、あるいは「快」と「不快」に属する単語を弁別することは、多くの人にとって容易であり、その反応時間は短く誤答もごく少ない。同様に、「花」と「快」の連合が強く、「虫」と「不快」の連合が強い人の場合は、「『花』または『快』」カテゴリーの単語と「『虫』または『不快』」カテゴリーの単語を弁別することも容易である。しかし、このような連合がつよいほど、「『花』または『不快』」カテゴリーと「『虫』または『快』」カテゴリーへの弁別課題を誤りなく迅速に実行することは困難になる。たとえば、<ごきぶり>という刺激が画面上に表示されると、「不快」概念が自動的に連想されるのに反して、「虫」と「快」を同じカテゴリーに分類しなければならないため、意識的な判断処理を注意深く実行する必要があるからである。

したがって、「花」と「快」、「虫」と「不快」の連合が強い参加者ほど、一致ブロック (「花」または「快」vs「虫」または「不快」) と不一致ブロック (「花」または「不快」vs「虫」または「快」) との反応時間差が大きくなると予想される。IAT は、このような一致ブロックと不一致ブロックにおける反応時間差を測定することで、概念間の連合の強さを相対的に推測しようとするものである。

実際の IAT は、表2のような7ブロックで構成されることが多く、第3・4ブロックが一致ブロック、第6・7ブロックが不一致ブロックである。Greenwald ら (1998) は当初、本試行 (ブロック4,7) の40個の反応時間のうち最初の2試行を除外し、第3試行以降の反応時間のうち300ms以下のものは

表 2 IAT の具体的な構成例

ブロック	試行数	左キー	右キー
1	20	花	虫
2	20	快	不快
3	20	花 + 快	虫 + 不快
4	40	花 + 快	虫 + 不快
5	20 または 40	虫	花
6	20	虫 + 快	花 + 不快
7	40	虫 + 快	花 + 不快

※実際には、カウンターバランスをとるために、半数の実験参加者には第 3・4 ブロックと第 6・7 ブロックの試行を入れ替えて実施する。

300ms に、3000ms 以上のものは 3000ms に置き換えた上で、それらに対数変換したものの平均値をブロック毎に算出し、その差を IAT スコアと呼び、連合の強さの指標とした。

しかし、この算出方法では、本来測定したいものとは別の、測定方法に依存する個人差（method-specific variance）を含むことが指摘され、Greenwald ら（2003）は新しいスコア算出方法として効果量の一種である D スコアを提案した。現在、D スコアが標準的な IAT スコア算出方法として広く使われている。

D スコア算出の詳細は割愛するが、スコア計算の対象となるのは第 3・4 ブロックおよび第 6・7 ブロックすべての試行の反応時間で、反応時間が 10000ms 以上のデータのみを分析から除外する。また、300ms 未満の反応時間が 10% 以上の実験参加者に対しては D スコアの算出は行わない。表 2 のような実験デザインの場合、D スコアが負でその絶対値が大きいほど、虫よりも花を好む潜在的傾向が強いと判断し、逆に D スコアの値が正の方向に大きくなるほど、虫をより好んでいることになる。

一般に、不慣れな課題に対して実験参加者のパフォーマンスは不安定で、試行を繰り返すうちに慣れが生じ、安定してくる。D スコアでは、練習ブロックや 1、2 回目の試行の反応時間も用いるなど、実験参加者のパフォーマンスが不安定な時期の結果を重視する算出方法となっている。

### 3. 拡散過程モデルによる IAT データの分析

拡散過程モデルでは、正答の場合の反応時間は初期値  $z$  から出発した拡散過程が閾値  $a$  に達するまでの時間に非決定時間  $t_0$  を加えたもの、エラー回答の反応

時間は拡散過程が閾値 0 に達するまでの時間に  $t_0$  を加えたものとしてモデル化される。このような仮定のもとで、正答および誤答の反応時間分布は、それぞれ、上述の 7 つのパラメータを用いて数学的に陽に表すことができるため、実験から得られた反応時間分布を理論分布に当てはめることによって、モデル・パラメータを推定することができる。詳しくは、例えば、Wagenmakers（2009）を参照のこと。

#### (1) 拡散過程モデルを適用するための IAT の実験デザイン

IAT の実験デザインには、間違っただけの回答に対して訂正を求め、刺激呈示から誤答が訂正されるまでの時間を反応時間として記録する場合と、訂正入力を求めず、誤答までの時間を記録する場合とがある。一般によく用いられるのは前者であるが、拡散過程モデルを用いたパラメータ推定を行う場合は、後者の実験デザインが望ましい。前者の実験デザインを用い、誤答までの反応時間を記録しない場合は、簡易推定法である EZ 法しか利用できない。

なお、訂正入力を求めず、誤答までの時間を記録する場合、D スコアとしては、 $D_3$  スコア～ $D_6$  スコアのいずれかを算出すればよい（Greenwald *et al.*, 2003）。

また、個人差測定を目的としてパラメータ推定を行う場合、第 4 および第 7 ブロックの試行数は可能な範囲で多く設定することが望ましい。

#### (2) 拡散過程モデルをどのように IAT データ分析に当てはめるか

個人差測定を目的として IAT データを分析する場合、一致ブロックおよび不一致ブロックに含まれる試行数は練習ブロックを含めてもせいぜい 100 個程度であると思われるので（標準的な IAT のデザインでは、60 個）、次のようにモデルの簡略化を行う（Klauer *et al.*, 2007）。

同一ブロック内では標的概念に属する刺激語と属性概念に属する刺激語が交互に表示されるが、拡散過程モデルに当てはめる場合にはこれら 2 種類の刺激を区別せず、また、単語の違いを区別しない。刺激の種類にかかわらず、刺激呈示に対する回答が「正答」と「誤答」のどちらであったかだけに着目し、一致ブロックと不一致ブロックそれぞれにおける実験参加者の情報





いと主張している。

IAT の一致ブロック (不一致ブロック) から得られるデータ数はそれほど多くないので、推定するパラメータは  $(v, a, t_0, s_{t0})$  の 4 つ、あるいは  $(v, a, t_0)$  の 3 つとするのが妥当であろう。より多くのパラメータを推定することも可能であるが、その場合、過剰適合 (オーバー・フィッティング) の問題がついてまわることになる (Ratcliff&Childers, 2015)。

では、一致ブロックと不一致ブロックで、共通の値として推定すべきパラメータはあるだろうか? たとえば、同じ被験者に対して、一致ブロックと不一致ブロックで非決定時間  $t_0$  は同じ値であると仮定して推定すべきであろうか? 先行研究 (Klauer *et al.*, 2007; van Ravenzwaaij *et al.*, 2011) では、不一致ブロックの方が情報収集率  $v$  は小さく、判断の慎重さ  $a$  と非決定時間  $t_0$  は大きくなることが示されており<sup>3</sup>、ブロック間で値を共通化すると、その他のパラメータの推定値に望ましくない影響を与えることが予想される。Voss ら (2015) においても、実験条件毎に、それぞれ別々にパラメータを推定することが推奨されている。

なお、II-1 で述べたように、拡散過程モデルにおいては、Wiener 過程のゆらぎの強さ  $s^2$  がスケーリングパラメータであるため、非決定時間  $t_0$  以外のパラメータの推定値は、Wiener 過程のゆらぎ強度  $s^2$  を基準とした値である。fast-dm では  $s^2 = 1$  を、EZ 法などでは  $s^2 = 0.01$  を用いているため、推定値の比較を行う場合は、スケール変換が必要になる。

## 2. どのパラメータ推定法を用いるべきか

拡散過程モデルに基づくパラメータ推定する場合、最尤法、カイ二乗法および Kolmogorov-Smirnov 法 (以下では、KS 法と表記) による推定が利用できる。また、正答の場合の反応時間の平均と標準偏差、および正答率だけから 3 つのパラメータ  $(v, a, t_0)$  の値を簡易的に算出する EZ 法がある。IAT データに対して、どの推定法を適用すればよいだろうか。

まず、II-3-(1) で述べたように、間違っただけに対して訂正を求め、正しく訂正されるまでの時間を反応時間として記録する場合は、EZ 法しか適用できない。また、EZ 法は正答率 100% の場合には解が得られないので、誤答がない場合にはパラメータを推定することができない。土居・川西 (2014) では、母親

258 名の子どもや養育に対する被害者意識を IAT によって測定データに EZ 法を適用したが、そのうちの 33% にあたる 84 名が一致ブロックと不一致ブロックのいずれかで正答率 100% であったため (多くは一致ブロックで)、IAT スコア ( $IAT_v$ ,  $IAT_a$ ,  $IAT_{t0}$ ) を推定できなかった。Rebar ら (2015) は、身体活動に対するシングル・カテゴリー IAT において、全体の 22%、23 名の正答率が 100% であったと報告した上で、正答率を 99.5% (これは、0.5 回の誤答に相当) に置き換えてパラメータを推定しているが、このような置き換えが妥当なものであるかどうかは不明である。

次に、間違っただけに対して訂正を求めず、誤答までの反応時間を記録した場合について述べる。IAT に組み込まれる試行数はそれほど多くないため、IAT データの分析には最尤法、KS 法および EZ 法が現実的に選択可能な推定法であろう。EZ 法以外は、正答率 100% の被験者にも適用可能である。

Voss ら (2013) は試行数と推定法との関係について表 3 のようにまとめている。試行数が 100 未満の場合には最尤法が第一選択肢であるが、最尤法は外れ値に敏感であり、とくに、短すぎる反応時間に対して非常に弱い。異常値が含まれる可能性があるなど、ロバストな推定が必要な場合は、100 未満のデータ数の場合でも KS 法を使うのが望ましい (Voss *et al.*, 2013; Ratcliff&Childers, 2015)<sup>4</sup>。

表 3 パラメータ推定における最適化基準の比較 (Voss *et al.*, 2013)

	最適化基準		
	最尤法	カイ二乗法	K-S 法
有効性	高い	低い	高い
頑健性	低い	高い	高い
計算スピード	速い	速い	遅い
必要な試行数 (およその目安)	少ない (N>40)	多い (N>500)	中程度 (N>100)

図 4 は、花 - 虫 IAT の一致ブロックの 60 個の反応時間からなる 5 組のデータセットに対して、 $(v, a, t_0, s_{t0})$  を推定し、150ms (誤答) と 200ms (正答) の混入データを意図的に加えることによって  $v$  と  $t_0$  の推定値がどのように変化するかをプロットしたものである。横軸が混入のないデータからの推定値、縦軸が 2 個の混入データを含めた場合の推定値である。図

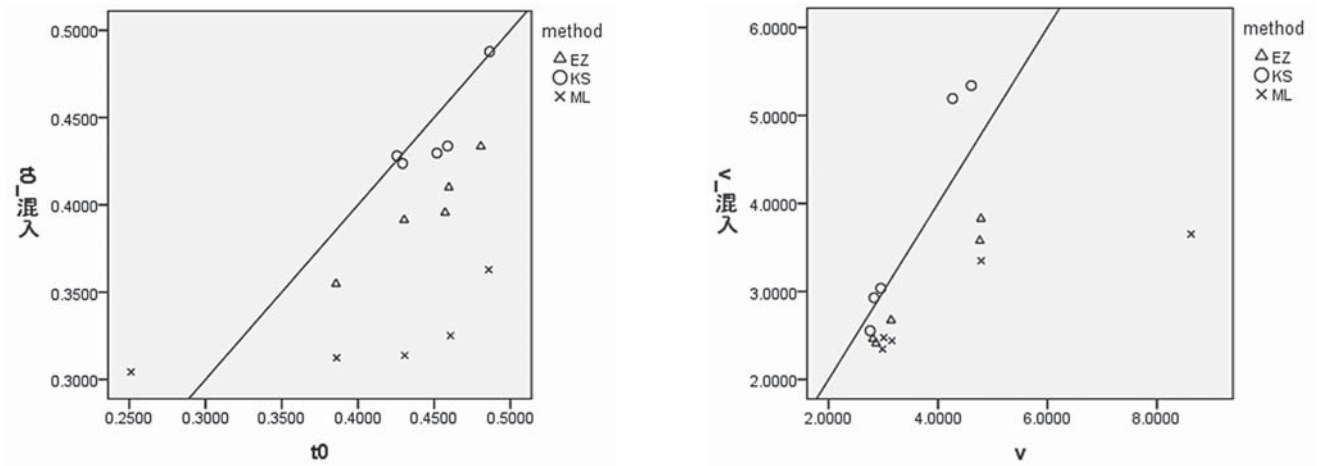


図4 短すぎるデータの混入に対する推定値の変化

花 - 虫 IAT の一致ブロックの 60 個の反応時間に 150ms と 200ms の混入データを混入させた。5 組のデータセットについてプロット。 $(v, a, t_0, s_{t0})$  を推定し、 $v$  と  $t_0$  について示している。

5 は、同じデータに対して、5000ms（正答）のデータを 1 つ混入させた場合である。KS 法がデータ混入にあまり影響を受けないのに対して、最尤法と EZ 法の推定値は強く影響を受けることが分かる。最尤法は特に短すぎるデータが混入することにより非決定時間  $t_0$  の推定値が非常に小さくなってしまい、帳尻を合わせるために情報収集率  $v$  の値が小さくなってしまふことが分かる。一方、EZ 法は反応時間の平均と標準偏差から算出する簡易推定法であるため、長すぎる反応時間が 1 つ混入するだけで  $t_0$  が負の値になってしまい、推定に失敗している。

EZ 法や最尤法を使うことで数十個のデータから相応の精度でパラメータを推定出来る可能性があるが、

そのためには、混入データや外れ値を事前に除去することが必須である。また、外れ値処理を十分に行うことが難しい場合は、EZ 法や最尤法は使うべきではない。KS 法を使うのが望ましい。

### 3. データの前処理

#### (1) 外れ値処理

上で述べたように、IAT データを拡散過程モデルに当てはめる場合には、外れ値処理が重要となる。Voss ら（2013）は、KS 法を使う場合には、遅すぎる外れ値と速すぎる外れ値を除外するために、反応時間の上限として 5000ms、下限として 200ms を設定する程度で十分なことが多いと述べている。これらの値は絶対

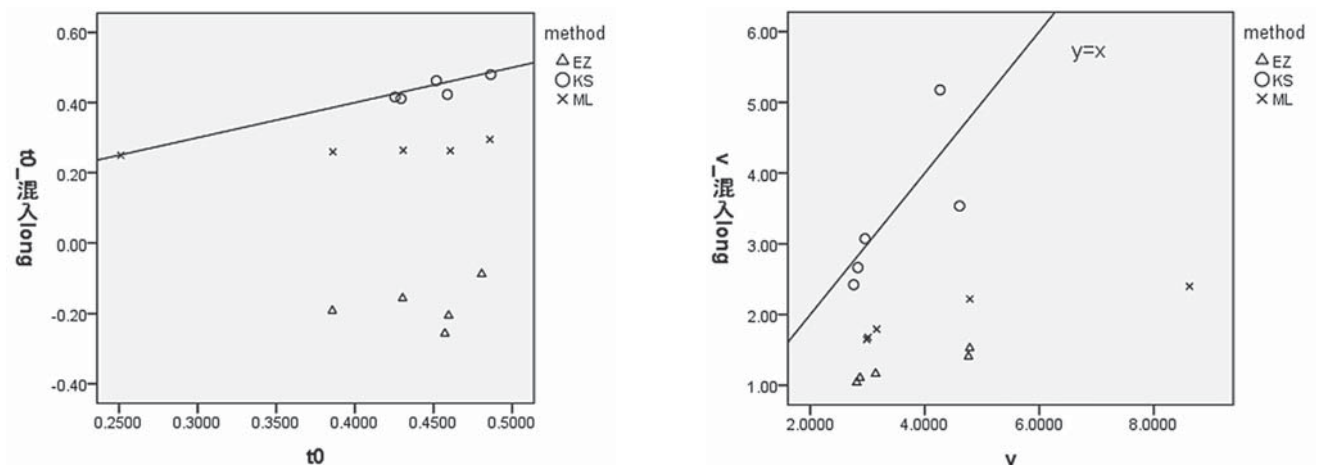


図5 長すぎるデータの混入に対する推定値の変化

花 - 虫 IAT の一致ブロックの 60 個の反応時間に 5000ms の混入データを混入させた。5 組のデータセットについてプロット。 $(v, a, t_0, s_{t0})$  を推定し、 $v$  と  $t_0$  について示している。

的なものではなく、たとえば van Ravenzwaaij ら (2012) では、下限を 275ms としている。

しかし、実際には、最尤法や EZ 法だけでなく KS 法を用いる場合でも、箱ひげ図に基づいた Tukey の基準に基づいて<sup>5</sup>実験参加者毎の反応時間分布をより厳密に検討することが多い。この基準は、実験参加者毎の反応時間分布の四分位範囲の値を  $h$  として、下限を (第 1 四分位点  $-1.5h$ )、上限を (第 3 四分位点  $+1.5h$ ) とする基準である<sup>6</sup>。Voss ら (2015) は、速すぎる応答を出来るだけ厳しく検出するために、反応時間を対数変換してから、上述の Tukey の基準を適用することを推奨している。

## (2) ウォーミングアップ効果をどう扱うか

通常の外れ値処理とは別に、試行開始時に見られるウォーミングアップ効果をどう扱うかという問題がある。図 6 に示すように、各ブロックの実験開始時、最初のいくつかの試行の反応時間はそれ以降と比べて明らかに長くなる。D スコアではすべての試行の反応時間からスコアを計算するが、拡散過程モデルは定常過程を想定したモデルであるので、最初のいくつかの試行から得られるデータは分析から除外するのが望ましい。とくに、外れ値処理を厳しく行わない場合は、注意する必要がある。

## 4. パラメータ推定の実行

筆者らは、Voss ら (2015) が GNU 一般公衆ライセンスに基づき公開している fast-dm-30 の C ソースコードを Windows 環境に移植し、グラフィカルユーザインタフェース (GUI) で気軽に利用できる分析パッケージを開発中である<sup>7</sup>。図 7 はプログラムの操作画面で、画面上で推定条件やデータを保存しているフォルダを指定するだけで、簡単にパラメータ推定を実行することができる。

以下では、この分析パッケージを用いて、花 - 虫 IAT データからパラメータ推定した結果を紹介する。分析に用いたデータは、女子大学生 78 名を対象に、表 2 の第 4 および第 7 ブロックの試行数を 100 とした花 - 虫 IAT を実施した結果である。一致ブロック (第 3・4 ブロック) と不一致ブロック (第 6・7 ブロック) に、それぞれ 120 試行を組み込んでいるが、Tukey の基準に従って外れ値処理を行った上で、さらに 200ms 未満のデータを削除した結果、約 13% のデータが外れ値として除外された。

$z=0.5$ 、 $s_v=0$ 、 $s_z=0$  とし、一致ブロック、不一致ブロック毎に 4 つのパラメータ ( $v$ ,  $a$ ,  $t_0$ ,  $s_{t0}$ ) を推定した結果を図 8 に示す ( $s_{t0}$  の推定結果は省略)。先行研究と同様、不一致ブロックの方が、慎重さのパラメータ  $a$  と非決定時間  $t_0$  は大きく、情報処理スピード  $v$  は小さくなっている。特に、情報処理スピード  $v$  で顕著な差がみられる。

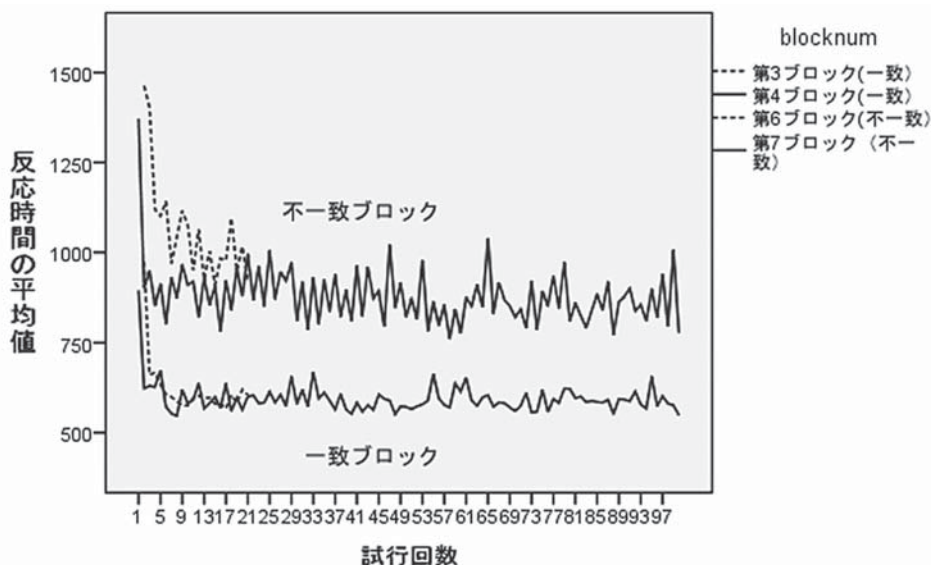


図 6 各ブロックにおける反応時間の変化。女子大生 78 名からの実験結果 (花 - 虫 IAT)



拡散モデルに基づくパラメータ推定プログラム

パラメータ値

precision3methodksmlcs

a推定する▼

zrset▼0.3<=0.5<0.7

v推定する▼

t0推定する▼

szrset▼0<=0<0.5

svset▼0<=0<2

st0推定する▼

dset▼-0.1<0<0.1

format

実験条件あり(stimulus RESPONSE TIME)

実験条件なし(RESPONSE TIME)

ファイル名

フォルダ名C:\kaken\FlowerInsect\FI\fastdm\_t15\DATA120Cut200ms\Fフォルダ選択

loadファイル名\*datファイル選択

saveファイル名\*\_ks.out , \_ml.out , \_ez.out

logファイル名20160913\_132359.log

コントロールファイル名experiment\_ksctl , \_mlctl

コントロールファイル作成⇒推定【fast-dm法】⇒累積確率分布【plot-cdf】正答率算出⇒グラフ化【累積確率分布図】⇒グラフ化【正答率散布図】

ファイル名	エラー内容	ファイルパス

リセット

図 7 パラメータ推定プログラムの操作画面

Figure 8 consists of three bar charts (a, b, c) showing the average values of parameters estimated by EZ, KS, and ML methods for one-block and non-one-block conditions. The y-axis represents the average value of the parameter, and the x-axis represents the method. Error bars indicate the 95% confidence interval.

Legend: ■ 一数ブロック (One-block), ▨ 不一数ブロック (Non-one-block)

(a) Parameters a, zr, and v. The y-axis ranges from 0.0 to 1.5. For each method, the non-one-block condition generally yields higher average values than the one-block condition.

(b) Parameters t0, szr, and sv. The y-axis ranges from 0.0 to 5.0. For each method, the non-one-block condition generally yields higher average values than the one-block condition.

(c) Parameter d. The y-axis ranges from 0.0 to 0.6. For each method, the non-one-block condition generally yields higher average values than the one-block condition.

図 8 EZ 法、KS 法、最尤 (ML) 法による推定値の平均と 95%信頼区間 (花 - 虫 IAT、N=78)



次に、各プロセス要素の IAT スコアである  $IAT_v$ 、 $IAT_a$ 、 $IAT_{t0}$  を算出し D スコア<sup>8</sup>との相関を算出した結果を表 4 に示す。EZ 法、KS 法、最尤法による推定値はほぼ同様の傾向を示しており、いずれの場合も、情報処理スピード  $v$  の IAT 成分が最も強く D スコアと相関している。これは、土居・川西 (2014) および Rebar ら (2015) と同様の結果であり、拡散過程モデルにおける概念間の連合の強さの指標は、 $IAT_v$  であると考えてよさそうである。

表 4 プロセス要素の IAT 成分と D スコアとの相関係数

	EZ 法	KS 法	最尤法
$IAT_v$	-.519**	-.493**	-.449**
$IAT_a$	.219	.175	.247*
$IAT_{t0}$	.418*	.272*	.313*

$N=78$ , \* $p<.05$ , \*\* $p<.01$

## 5. 推定結果の検討

では、拡散過程モデルの実データへの当てはまりは、どのように評価したらよいだろうか。KS 法の場合は  $p$  値、最尤法の場合は最大対数尤度が出力されるが、EZ 法はモーメント法による簡易推定であるため適合度指標は得られない。また、IAT の場合、十分な試行数を用意することが難しいので、KS 法で出力される  $p$  値による評価は甘くなる傾向にある。

Voss ら (2013, 2015) は、予測累積分布関数 (CDF) と実データを同じグラフに重ね書きすることで実験参加者毎・実験条件毎の当てはまりを評価し、実験参加

者全体に対する当てはまりは、正答率や反応時間分布の四分位点について横軸を実データ、縦軸を予測値とする散布図によって評価する方法を提案している。

図 9 は、正答率についての散布図である。横軸は実験参加者毎の正答率、縦軸は KS 法と最尤法からの予測正答率で、78 名の一致ブロックと不一致ブロックでの値をそれぞれプロットしている。(a) は 120 試行、(b) は 60 試行 (第 3 (6) ブロック 20 試行と、第 4 (7) ブロックの最初の 40 試行) から推定した結果である。EZ 法は、実データと等しくなるようにパラメータを算出する方法であるため、予測正答率は実データと同じ値になる。

120 試行からの推定の場合、最尤法による予測正答率はほとんどのデータで実データとほぼ一致するが、一部のデータ (データ ID: ICA012708, ICA012602) の予測値は実データから大きく外れている。一方、KS 法の場合は、多少の誤差が伴うものの、比較的安定した結果を示すことが分かる。図 11 (a) は、ICA012708 に対する反応時間の累積分布関数の予測値を実データに当てはめた結果である。最尤法が分布の当てはめに失敗していることがわかる。

60 試行から推定した場合も、やはり最尤法は高い精度で実正答率を再現している。また、120 試行の場合に見られるような失敗例も見当たらない。しかし、実はいくつかのデータに対して過剰適合が起こっている。図 10 は、 $v$  と  $a$  の推定値について、横軸を EZ 法からの推定値、縦軸を KS 法と最尤法からの推定値として作成した散布図である。各パラメータの真の値

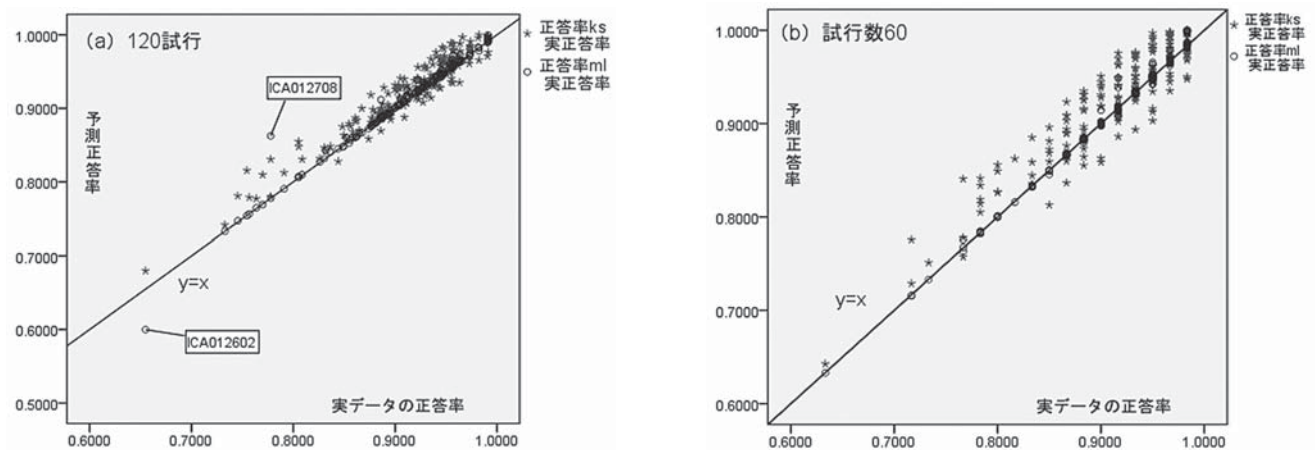


図 9 正答率の散布図 (a) 120 試行からの推定結果 (b) 60 試行からの推定結果  
横軸：実データ、縦軸：予測正答率、 $N=78$

は分からないので、それと比較することは出来ないが、いくつかのデータにおいて、最尤法が他の推定法や他のデータと大きく異なる値を出力していることがわかる。図 11 (b) は、最尤法による  $v$  の推定値が非常に大きくなっているデータ CA012701 に対する反応時間分布の当てはまり状況である。最尤法は EZ 法や KS 法よりも高い精度で累積分布関数を再現していることから、データ数が少ない場合に起こりがちな過剰適合が生じている可能性が高い。図 10 (a) (b) で見られる外れ値 CA012205、CA012615、CA12702、ICA012709 においても、最尤法は実データの分布関数を高い精度で再現していることから、同様の問題が生じていると考えられる。

#### IV おわりに

今回の実験データでは、120 試行から推定した各プロセス要素の IAT スコア ( $IAT_v$ 、 $IAT_a$ 、 $IAT_{t0}$ ) の値は、EZ 法、KS 法、最尤法のどれを用いても大きな違いはなく、推定法間の相関は 0.8 ~ 0.9 程度であった。しかし、60 試行からの推定値は、EZ 法と KS 法との相関が 0.85 前後であるのに対して、最尤法とその他の推定法との相関は、 $IAT_v$  は 0.5 程度、 $IAT_a$  は 0.3 程度であった。最尤法において過剰適合が発生し、 $a$  や  $v$  の推定に失敗しているのが原因であろう。

したがって、標準的な IAT から得られたデータを拡散過程モデルで分析する場合には、入念に外れ値処理をした上で EZ 法を用いるか、あるいは KS 法を用いるのが良さそうである。また、最尤法を用いる場合には、過剰適合にも十分な注意を払う必要がある。

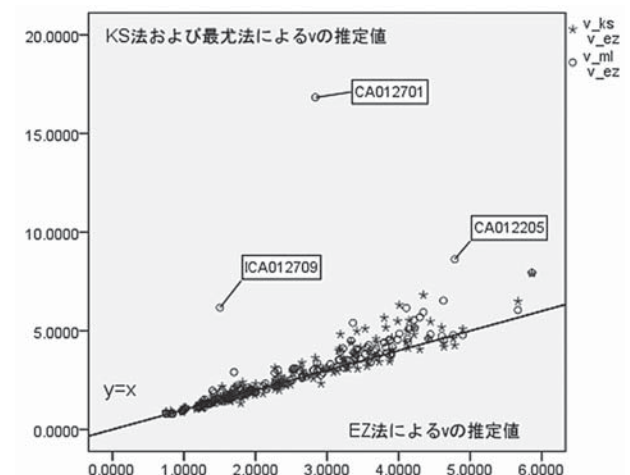
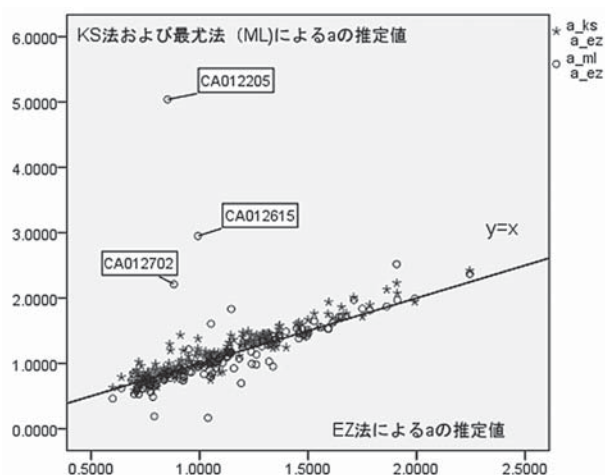


図 10 パラメータ推定値の比較 (60 試行からの推定結果) (a) 情報収集率  $v$ 、(b) 慎重さ  $a$   
横軸：EZ 法の推定値、縦軸：KS 法と最尤 (ML) 法の推定値

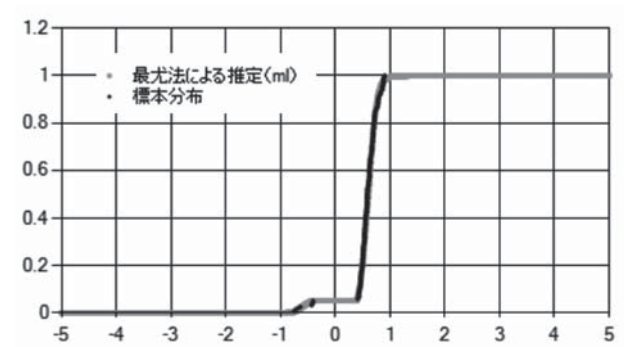
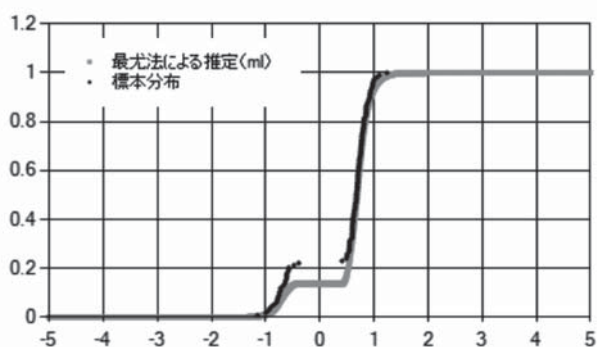


図 11 累積分布関数の推定結果

(a) ICA012708 の結果 (120 試行の場合) (b) CA12701 の結果 (60 試行の場合)

なお、今回、推定結果の当てはまりを確認するツールとして正答率の散布図を用いたが、この方法ではEZ法の推定結果を評価することが出来ない。筆者らが開発中の分析パッケージには、正答率に加えて、第1四分位点、第2四分位点、第3四分位点の散布図を出力する機能を追加する予定である。

## 注

- 1 スケーリングパラメータであるため任意の値を設定できる。 $s^2 = 0.01$ 、または、 $s^2 = 1$ と設定されることが多い。
- 2 fast-dm30では、8個目のパラメータが追加されている。
- 3 一般的に連合が強いと想定される概念の組み合わせを「一致ブロック」、それとは逆の組み合わせを「不一致ブロック」としている。
- 4 Ratcliff&Childers (2015)では、より適切な方法として階層的拡散過程モデルを挙げている。
- 5 反応時間分布は正に歪んだ分布なので、標準偏差を基準とする方法はあまり適切ではない。
- 6 統計解析ソフトウェアSPSSで箱ひげ図を出力すると、 $1.5h$ を基準にした外れ値は白丸(○)で、 $3h$ を基準としたさらに極端な値はアスタリスク(\*)で、表示される。
- 7 この分析パッケージと使い方の詳細は、下記Webサイトで公開予定である。  
<http://www.koka.ac.jp/DFmodel>
- 8 Dスコアの算出には、誤答に対して600msのペナルティを課す $D_4$ スコアを用いた (Greenwald *et al.*, 2003)。

## 引用・参考文献

- 土居・川西 (2012) 拡散モデルに基づく潜在的連合テストデータの分析, 京都光華女子大学研究紀要, 50, 111-122.
- 土居・川西 (2014) Ratcliffの拡散モデルに基づくIATデータの分析, 日本心理学会第78回大会発表論文集, 2PM-2-004.
- 土居・川西 (2015) Ratcliffの拡散モデルに基づく反応時間分析ソフトウェアの開発-Voss & Vossの

fast-dm-30をベースに, 日本心理学会第79回大会発表論文集, 1EV-062.

- Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. (1998) Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464-1480.
- Greenwald A. G., Nosek B. A. and Banaji, M. R. (2003) Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197-216.
- Klauer, K. C., Voss, A., Schmitz, F. & Teige-Mocigemba, S. (2007) Process-components of the Implicit Association Test: A diffusion model analysis. *Journal of Personality and Social Psychology*, 93, 353-368.
- Ratcliff, R. (1978) A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Ratcliff, R. & Childers, R. (2015) Individual differences and fitting methods for the two-choice diffusion model of decision making, *Decision*, 2 (4), 237-279.
- Ratcliff, R. & Tuerlinckx, F. (2002) Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9, 438-481.
- Rebar, A.L., Ram, N. & Controy, D.E. (2015) Using the EZ-diffusion model to score a single-category implicit association test of physical activity. *Psychology of Sport and Exercise*, 16 (3), 96-105.
- Vandekerckhove, J. & Tuerlinckx, F., (2008) Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, 40, 61-72.
- Vandekerckhove, J., Tuerlinckx, F. & Lee, M. D. (2011) Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 16, 44-62.
- van Ravenzwaaij, D. & Oberauer, K. (2009) How to use the diffusion model: Parameter recovery of

- three methods: EZ, fast-dm, and DMAT. *Journal of Mathematical Psychology*, 53 (6) :463-473.
- van Ravenzwaaij, D., van der Mass, H. & Wagenmaker, E. -J. (2011) Does the name-race Implicit Association Test measure racial prejudice? *Experimental Psychology*, 58, 271-277.
- Voss, A., Nagler, M., & Lerche, V. (2013) Diffusion Models in Experimental Psychology: A Practical Introduction. *Experimental Psychology*, 60, 385-402.
- Voss, A. & Voss, J. (2007) Fast-dm: A free program for efficient diffusion model analysis. *Behavioral Research Methods*, 39, 767-775.
- Voss, A., Voss, J. & Lerche, V. (2015) Assessing Cognitive Processes with Diffusion Model Analyses: A Tutorial based on fast-dm-30. *Frontiers in Psychology*, 6:336.
- Wagenmakers E.-J. (2009) Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, 21 (5), 641-671.
- Wagenmakers, E.-J., van der Maas, H. L. & Grasman, R. P. (2007) An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14, 3-22.
- Wiecki, T.V., Sofer, I. & Frank, M.J. (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python, *Frontiers in Neuroinformatics*, 7:14, DOI: 10.3389/fninf.2013.00014