

反応時間の確率分布推定

谷津 貴久

Estimating the Probability Distribution of Reaction Times

Takahisa TANITSU

要 約

反応時間の確率分布モデルをデータに適用する際、複数のモデルを比較選択する方法としてAIC (Akaike information criterion) の利用が挙げられる。しかしAICによって常に真のモデルを選択することができるのかという疑問が生じる。本稿では、対数正規分布・ガンマ分布・複合指数分布・指数正規合成分布・指数ガンマ合成分布・ガンベル分布・ワイブル分布・ワルド分布の8種の確率分布に対して、AICにより真の分布を正しく選択できるのかどうかを計算機シミュレーションにより検証した。その結果、複合指数分布と指数正規合成分布が真の分布である場合にはAICによって正しく選択することができず、試行数が100以下ではどの分布も正しく選択することが困難であることが分かった。また、真の分布にかかわらずガンマ分布が選択される率が高いという結果も得た。以上より、AICのみによって反応時間の確率分布モデルを選択することには慎重になるべきであると結論した。

キーワード：反応時間，確率分布，AIC，シミュレーション

1. 確率分布としての反応時間

心理学が測定する指標として、正答率と共によく用いられているのが反応時間である。反応時間は一般に正歪分布するといわれ (宮埜, 1993), 図1のような形の確率密度関数で近似できる。通常この反応時間の分布は対数正規分布 (lognormal distribution) であるとみなされ, 分散分析などデータに正規分布が仮定される分析を適用する際には前処理として原データの対数変換が行われる。この典拠として, 上田・中本・谷津・椎名 (1999) は Woodworth and Schlosberg (1954) を挙げている。しかし図1の形を対数正規分布とみなす理論的必然性はなく, 同様の形をしているガンマ分布 (gamma distribution) やその他の分布でも構わないはず

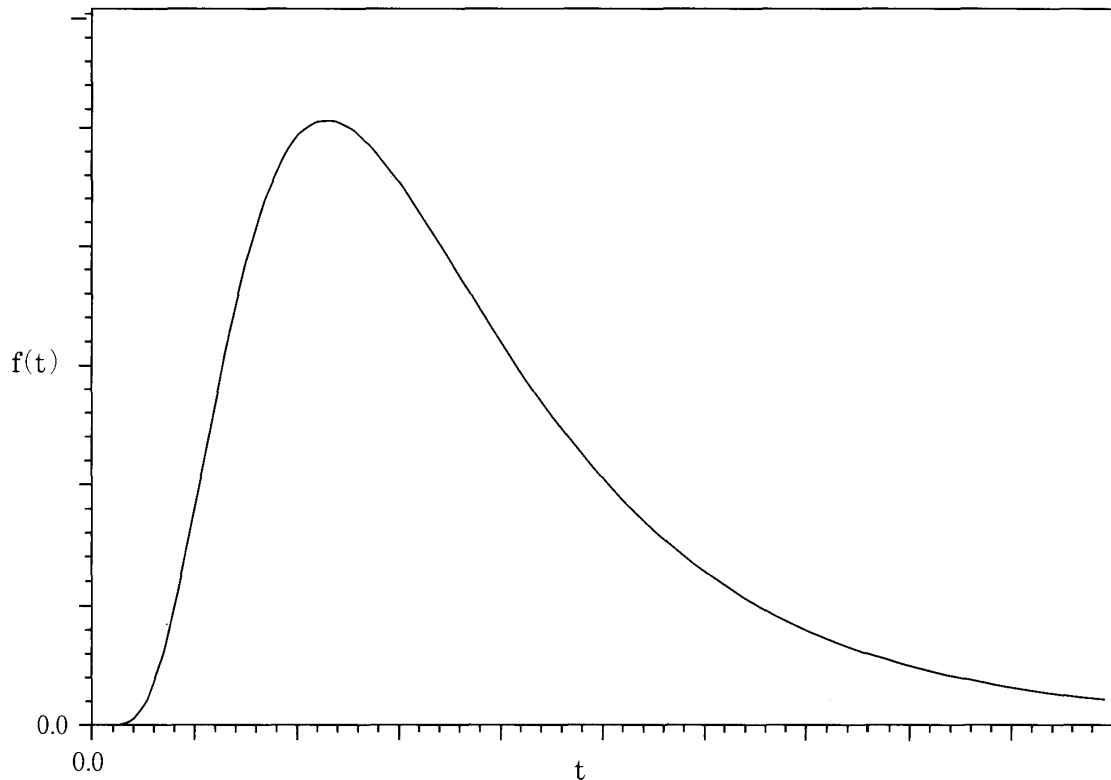


図1 仮想的な反応時間の確率分布

である。

一方、刺激が与えられてから反応が生起するまでには様々な心理学的・生理学的処理過程が介在し、それぞれに処理時間があると考えられる。心理学実験において課題の種類により反応時間に差が生じたならば、その中の全ての処理過程が課題の種類の影響を受けたと考えるよりも、影響を受けて処理時間に変化が生じた過程と影響を受けない過程とがあると考えるのが自然である。反応時間はそれら個々の処理時間から構成されるデータと考えることができる。これを逆に考えれば、実験により得られた反応時間を個々の処理時間に分解することができるはずである。刺激から反応までの処理過程はいくつにでも細かく分けて考えることができるが、心理学の実験精度を考慮すると大きく2つの処理過程に分解するのが妥当であろう。また、2つの処理過程を経て得られた反応時間の分布は図1のような形をしていなければならないという制約がある。

これらの点を考慮してこれまで図1のような形をする様々な確率分布モデルが提唱されてきた (Luce, 1991; 上田・中本・谷津・椎名, 1999, を参照)。それらモデルの検証は、実験か

ら得られた反応時間データへのモデルのあてはまりの良さを見ることで行われる。モデルのあてはめに最尤法を用いれば、その時に得られる AIC (Akaike information criterion, 坂元・石黒・北川, 1983, を参照) などの指標によって複数モデル間であてはまりの良さの比較が行える。

ここで問題となるのは、AIC のような指標によってあてはまりが良いと判定されたモデルが真の確率分布の正しい推定となっているのかという点である。無限のサンプルが得られれば真のモデル以外のあてはまりは悪くなり、AIC などの指標で正しくモデルの選択ができる。しかし実験の試行数は限られており、限られた数の反応時間データしか得られない。例として対数正規分布に従う擬似乱数を 10 個生成したときのヒストグラムを図 2 に示す。図 2 から明らかなように、試行数が少ない場合には真の確率分布（この場合は対数正規分布）に近い形のきれいな分布は得られない。したがってそのようなデータに確率分布モデルをあてはめたとしても真のモデルが最もよくあてはまるとは限らず、真のモデルとは異なるモデルが最良と誤判定される危険性があると考えられる。

本稿では、反応時間の確率分布モデルを推定する際に AIC による判定が有効なのかどうか、有効ならばどの程度の試行数があれば正しい推定を行えるのかという問題を扱う。以下に代表

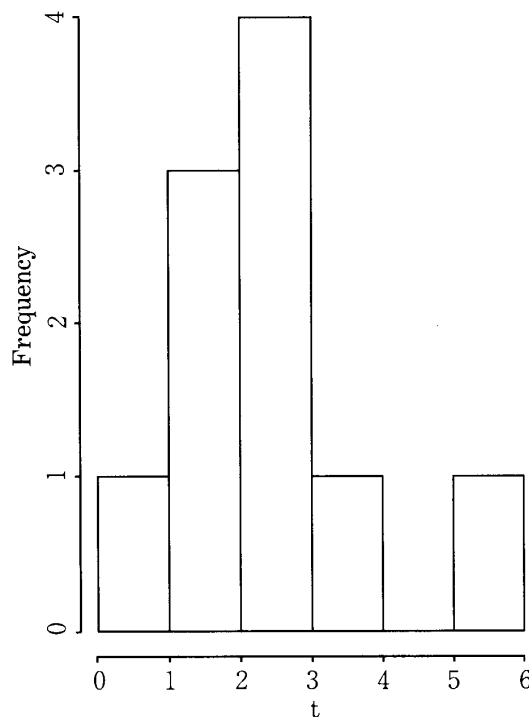


図 2 対数正規分布（平均 2.0, 分散 1.0）に従う擬似乱数を 10 個生成したときのヒストグラム

的な反応時間の確率分布モデルを取り上げ、試行数の多少が確率分布モデルの推定に与える影響を計算機シミュレーションによって検討する。

1.1 代表的な反応時間の確率分布モデル

本稿で検討する確率分布モデルは大きく4種に分けられる。まず処理過程についての仮定が何も無い単純な確率分布、次に直列につながった2種の処理過程を記述する直列過程、並列する2種の処理過程を記述する並列過程、そしてある閾値に達すると反応が生起する処理過程を記述する蓄積過程である。

なお、特に断わらない限り、以下に挙げるそれぞれの確率密度関数は基本的に蓑谷 (1998)、Luce (1991) に取り上げられた関数である。これらの関数は基本的に全て図1の形をしている。本稿では各確率分布モデルの理論的背景や適用実験については原則として触れない。それらはLuce (1991)、上田・中本・谷津・椎名 (1999) に詳しく述べられている。

単純な確率分布

反応時間のモデルとして提唱されている確率分布モデルは、処理過程に関する何らかの心理学的仮定を含んでいる。しかし本稿では、処理過程について何ら言及しない単純な確率分布も比較のために取り上げる。

対数正規分布 (lognormal distribution) Woodworth and Schlosberg (1954) 以来、反応時間の確率分布モデルとして対数正規分布が暗黙裡に仮定されてきた。対数正規分布の確率密度関数は次式で表わされる。

$$f(t) = \frac{1}{\sqrt{2\pi} \sigma t} \exp \left[-\frac{(\log t - \mu)^2}{2\sigma^2} \right], \quad (1)$$

ただし \log は自然対数を表わす。

ガンマ分布 (gamma distribution) パラメータの等しい指数分布 (exponential distribution) を α 個合成 (たたみ込み) したのが次式のガンマ分布である。この点からガンマ分布を次節の直列過程とみなすこともできるが、等しいパラメータという制約と合成する指数分布の個数に制限がないという点から本稿では直列過程とはみなさず、単純な確率分布として扱う。

$$f(t) = \frac{t^{\alpha-1} \exp(-t/\beta)}{\beta^\alpha \Gamma(\alpha)}, \quad (2)$$

ただし $\alpha > 0$, $\beta > 0$ で, $\Gamma(\cdot)$ はガンマ関数である。

直列過程

一方の処理過程終了後に次の処理過程が始まるという, 2つの処理過程の単純な結合が直列過程である。直列過程では反応時間は2つの過程の処理時間を足したものになる。一般に指数分布ともうひとつの分布の合成すなわちたたみ込みによって得られる分布が, 反応時間の確率分布モデルとして提唱されている。

複合指数分布 (ex-exponential distribution) ガンマ分布がパラメータの等しい指数分布の合成であったのに対し, パラメータの異なる2つの指数分布を合成したものが複合指数分布である (佐々木・藤田, 1993)。

$$f(t) = \frac{\alpha\beta}{\beta - \alpha} [\exp(-\alpha t) - \exp(-\beta t)], \quad (3)$$

ただし $\alpha > 0$, $\beta > 0$ である。

指数正規合成分布 (ex-Gaussian distribution) 指数分布と正規分布 (normal distribution, Gaussian distribution) との合成によって得られる分布である。指数正規合成分布の確率密度関数は次式のようになる。

$$f(t) = \lambda \exp(-\lambda t) \exp\left(\lambda\mu + \frac{1}{2}\lambda^2\sigma^2\right) \Phi\left(\frac{t-\mu-\lambda\sigma^2}{\sigma}\right), \quad (4)$$

ここで $\lambda > 0$, また $\Phi(\cdot)$ は標準正規累積分布関数である。

正規分布を取り込んでいるため, このモデルでは $-\infty < t < \infty$ となる。したがって厳密には反応時間のモデルとしては適さない。Luce (1991) は $\mu > 3\sigma$ の制約を付けることによって $t < 0$ になる可能性を小さくすることを提案している。

指数ガンマ合成分布 (ex-gamma distribution) 指数分布とガンマ分布との合成分布である。

Suppes, Cohen, Laddaga, Anliker, and Floyd (1983) によって提唱された。確率密度関数は次式のとおりである。

$$f(t) = \frac{\alpha}{\lambda_1} \exp\left(-\frac{t}{\lambda_1}\right) + \frac{(1-\alpha)t}{\lambda_2^2} \exp\left(-\frac{t}{\lambda_2}\right). \quad (5)$$

並列過程

並列する2つの処理過程を考えているのが並列過程のモデルである。並列過程では、双方の処理時間の最小値が反応時間となるモデルと最大値が反応時間となるモデルとを考えることができる。すなわち、前者は並列する処理過程のどちらかが終われば反応が生起し、後者は双方の処理が終わるまで待ってから反応が生起するモデルである。以下に最大値の場合の確率分布を挙げる。最小値との関係については上田・中本・谷津・椎名(1999)に述べられている。

ガンベル分布 (Gumbel distribution) 二重指数分布 (double exponential distribution) あるいは極値分布タイプ I (extreme value distribution of type I) とも呼ばれる。ガンベル分布の確率密度関数は次式で表わされる。

$$f(t) = \frac{1}{\theta} \exp\left(-\frac{t-\mu}{\theta}\right) \exp\left[-\exp\left(-\frac{t-\mu}{\theta}\right)\right], \quad (6)$$

ここで $-\infty < t < \infty$ なので、反応時間の確率分布モデルとしては厳密に言えば適切ではない。

ワイブル分布 (Weibull distribution) 極値分布タイプ III (extreme value distribution of type III) とも呼ばれる。確率密度関数は次式で表わされる。

$$f(t) = \frac{\gamma}{\phi} \left(\frac{t}{\phi}\right)^{\gamma-1} \exp\left[-\left(\frac{t}{\phi}\right)^\gamma\right]. \quad (7)$$

蓄積過程

刺激を受容することによって心的な何かが蓄積され、それが閾値を越えることによって反応が生起すると考えるモデルである。したがって、このモデルは反応時間を分解することができる直列過程・並列過程とは質的に異なる。

ワルド分布 (Wald distribution) 逆ガウス分布 (inverse Gaussian distribution) とも呼ばれ、

確率密度関数は次式で表わされる。

$$f(t) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left[-\frac{\lambda(t - \mu)^2}{2\mu^2 t}\right]. \quad (8)$$

反応時間のモデルとしては、 $\lambda = (C/s)^2$ 、 $\mu = C/m$ と置き C を閾値とみなす (Luce, 1991; 上田・中本・椎名・谷津, 1999)。このモデルを適用することにより閾値を推定することができる。

2. シミュレーション

複数の反応時間の確率分布モデルを実際のデータにあてはめた際に、AICによって真の分布が正しく選択できるか、またどの程度の試行数が必要になるのかを計算機シミュレーションによって検討した。

シミュレーションは、前節に挙げた8種類の確率分布に従う擬似乱数を計算機により生成して反応時間データとみなし、それぞれに8種の確率密度関数をあてはめることによって行った。AICによって真のモデルが正しく選択できるならば、ある確率分布に従う擬似乱数に対して8種類のモデルをあてはめた場合、擬似乱数の生成に用いた確率分布モデルがAICによって選択されるはずである。さらに擬似乱数の生成数（実験での試行数にあたる）が少ないと、AICによって正しく選択できる可能性が低くなると考えられる。

擬似乱数の生成 擬似乱数の生成に用いた各分布のパラメータを表1に示す。このパラメータでほとんどの分布が期待値（平均値）2.0，分散1.0となる。ただし複合指数分布と指数ガンマ

表1 擬似乱数の生成に用いたパラメータ

確率分布	パラメータ
対数正規分布	$\mu = 0.58, \sigma = 0.47$
ガンマ分布	$\alpha = 4, \beta = 0.5$
複合指数分布	$\alpha = 1.34, \beta = 1.5$
指数正規合成分布	$\lambda = 1.2, \mu = 1.17, \sigma = 0.55$
指数ガンマ合成分布	$\alpha = 0.1, \lambda_1 = 3, \lambda_2 = 0.9$
ガンベル分布	$\theta = 0.78, \mu = 1.55$
ワイブル分布	$\gamma = 2.10, \phi = 2.26$
ワルド分布	$\mu = 2, \lambda = 8$

合成分布はこの期待値と分散にすることが不可能なため、期待値がそれぞれ 1.41 と 2.1, 分散が 1.0 と 0.93 になるパラメータに設定した。

このパラメータに基づいて、8種の各分布に従う擬似乱数を生成した。各分布ごとに1組あたりの擬似乱数の生成数を 10, 50, 100, 1000 の4種類設け、それぞれ 300組ずつ生成した。各組が実験における被験者に対応し、生成数が試行数に対応する。

確率分布モデルのあてはめ 生成した擬似乱数を反応時間データとみなし、1組ごとに8種の確率密度関数をあてはめた。パラメータの推定には最尤法 (maximum likelihood method) を用い、対数尤度と AIC を算出した。対数尤度関数の最大化には統計解析ソフトウェア R (<http://www.ci.tuwien.ac.at/R/>) の nlm 関数を用い、推定の初期値には表1の値を用いた。

3. 結 果

各分布に従う擬似乱数について、8種の確率密度関数をあてはめた結果 AIC が最小となったモデルを求めた。その中で擬似乱数の生成に用いた分布と同じモデルが選択された割合、すなわち真の分布を正しく選択できた割合を表2に示す。

表2よりまず明らかなのは、複合指数分布と指数正規合成分布は、自ら生成した擬似乱数にあてはめを行っても AIC で比較する限り他の分布より良いモデルとして選択される割合が非常に少ないという点である。つまり真のモデル以外が AIC で最良と選択されている。これは擬似乱数の生成数が 1000 まで増えても改善されない。

擬似乱数の生成数と真の分布の推定との関係について、上記2種以外の分布は擬似乱数の生

表2 真の確率分布が AIC により正しく選択された割合

単位：%

確 率 分 布	擬似乱数の生成数			
	10	50	100	1000
対数正規分布	0.0	1.0	3.3	58.3
ガンマ分布	0.0	4.7	32.7	95.3
複合指数分布	0.0	0.0	1.0	8.3
指数正規合成分布	0.0	0.0	0.0	0.0
指数ガンマ合成分布	7.3	39.3	52.7	82.3
ガンベル分布	1.0	11.7	28.7	86.7
ワイブル分布	17.0	54.7	68.7	99.7
ワールド分布	0.0	1.0	6.3	62.0

成数が1000になると真のモデルが選択される割合が増加した。ただしガンマ分布・指数ガンマ合成分布・ガンベル分布・ワイブル分布が生成数1000において80%以上の選択率になったのに対し、対数正規分布とワルド分布は60%程度にしかない。またいずれの分布も生成数が50では真のモデルが最良と選択される割合が低く、50%を越えているのはワイブル分布のみである。生成数が10ではほとんどのモデルで0%となる。つまり真のモデル以外が選択されている。

次に、AICで最良と選択されたモデルを生成数(10, 50, 100, 1000)別に集計した。それぞれの生成数ごとに8種の分布に従う擬似乱数が全て含まれているにもかかわらず、どの生成数においてもガンマ分布が最良と判定された割合が最も多かった。AICによりガンマ分布が選択された割合を生成数別に表わしたのが表3である。生成数が10の場合には、2400(8種×300組)のデータのうち半数以上においてガンマ分布が選択されている。生成数が増えるに従ってその割合は減少するが、生成数1000においても高い割合を示している。

表3 AICによりガンマ分布が選択された割合

単位：%

擬似乱数の生成数	ガンマ分布が選択された割合
10	52.3
50	42.9
100	37.1
1000	27.6

4. 考 察

反応時間の確率分布モデルを検証する実験において、AICにより真の確率分布モデルを選択することができるのか、またそれにはどの程度の試行数が要求されるのかという問題を本稿では扱った。以下にそれぞれの点について述べる。

AICによる真のモデルの選択について 計算機シミュレーションの結果を見ると、複合指数分布と指数正規合成分布が真の分布である場合には、AICによってそれらのモデルが選択される可能性が極めて低い。これら2種の分布のうち指数正規合成分布はパラメータ数が3であり、他のほとんどの分布よりもパラメータがひとつ多い(表1)。AICはモデルのパラメータ数も考慮する指標であることを考えれば、指数正規合成分布がAICで選択されにくいのは他のモ

デルよりパラメータが多いためという可能性がある。しかし同様に AIC が最小となりにくい複合指数分布はパラメータ数が 2 であり，また指数正規合成分布と同様にパラメータ数が 3 である指数ガンマ合成分布は AIC で最良と判定される割合が高い。よって今回のシミュレーション結果に関してはモデルのパラメータ数の影響は低いと考えられる。以上より，仮にこの 2 つの分布のいずれかが真の分布である場合には，AIC の比較のみによってそれらを推定することが困難であるといえる。

実際に反応時間の確率分布を求めた研究はいくつかあるが，多数の確率分布モデルをあてはめた研究として，椎名・上田・中本・森口・谷津（1999），上田・椎名・中本・谷津・森口（1999）のものがある。彼らは 1 人あたり 40 試行の単純反応時間データを 331 人分収集し確率分布モデルをあてはめた。彼らが扱った確率分布は本稿で取り上げた分布とほぼ同一であったが指数ガンマ合成分布はなく，代わりに正規分布を含めていた。データへの確率分布のあてはまりを AIC により比較した結果，複合指数分布が最良となることはなかった。しかしながら，本稿でのシミュレーション結果によると複合指数分布はそもそも AIC で最良と判定されにくい。そのため彼らが得た結果をもって複合指数分布を単純反応時間の確率分布モデルとして棄却することはできなくなる。

また椎名・上田・中本・森口・谷津（1999），上田・椎名・中本・谷津・森口（1999）では，331 人分のデータのうち 77 のデータで指数正規合成分布が AIC により最良と判定されている。今回のシミュレーション結果によれば，指数正規合成分布が真の分布である場合に AIC で正しく選択される可能性は極めて低い。この矛盾する結果を解釈することは今回の結果のみでは難しいため，今後さらに詳細な研究が必要である。

試行数とモデル選択の関係について 今回検討した 8 種の確率分布モデルは，生成数が 100 以下では真の分布が AIC によって選択される可能性が低いことが分かった。この点と関連して，生成数が 10 や 50 など少ない場合に AIC によるモデル選択を行うと，真の分布にかかわらずガンマ分布が選択される可能性の高いことが分かった。したがって，試行数の少ない実験によって得られた反応時間データに確率分布モデルのあてはめを行った際に，AIC によってガンマ分布が最良と判定されてもそれが真の分布であるとは断定できないことになる。

以上で本稿での主な論点について考察した。これらの解釈に対し，いずれも対数尤度関数の最大化が不完全であったために今回のような結果が生じたと反論することも可能である。つまり，パラメータ推定にまだ改良の余地があるのではないかという可能性である。そのような可能性を根本的に排除することはできない。しかしながら，パラメータ推定の初期値には擬似乱

数の生成パラメータを与えており、良い初期値であるといえる。したがってそのような反論が成立する可能性は低いと考えられる。

また、本稿では真の確率分布の存在を仮定し、反応時間の確率分布モデルについての研究は真の確率分布を推定することであるという考え方を採っている。しかし真のモデルの推定という考え方そのものが誤りであり、データにうまくあてはまる記述モデルが存在すればよいという考え方も成り立つ。その考え方を採用すれば、本稿のシミュレーション結果から反応時間の確率分布モデルとして最も妥当なのはガンマ分布であると結論することもできる。

当面の結論として、反応時間の確率分布モデルをデータにあてはめる際には、理論的裏付けなしにAICによって機械的にモデルの優劣を決めることには慎重になるべきである。理想的には、AICなどの指標を使わずに理論のみでモデルを選択できることが望ましい。そのために、反応時間の確率分布に関する理論的研究のさらなる進展が望まれる。

引用文献

- Luce, R. D. 1991. *Response Times. Their Role in Inferring Elementary Mental Organization* (Paperback edition). New York: Oxford University Press.
- 蓑谷千風彦 1998. 『すぐに役立つ統計分布』東京図書.
- 宮埜壽夫 1993. 『心理学のためのデータ解析法』培風館.
- 坂元慶行・石黒真木夫・北川源四郎 1983. 『情報量統計学』共立出版.
- 佐々木久子・藤田恵璽 1993. 「手先動作の熟達過程における所要時間分析」, 『人間工学』, **30**, pp. 179–186.
- 椎名乾平・上田卓司・中本敬子・森口果優・谷津貴久 1999. 「反応時間の分布とパラメーター推定 (1)」, 『日本心理学会第63回大会発表論文集』, p. 428.
- Suppes, P., Cohen, M., Laddaga, R., Anliker, J., & Floyd, R. 1983. “A procedural theory of eye movement in doing arithmetics”, *Journal of Mathematical Psychology*, **27**, pp. 341–369.
- 上田卓司・中本敬子・谷津貴久・椎名乾平 1999. 「反応時間の確率分布」, 『早稲田大学教育学部学術研究—教育心理学編—』, **47**, pp. 25–36.
- 上田卓司・椎名乾平・中本敬子・谷津貴久・森口果優 1999. 「反応時間の分布とパラメーター推定 (2)」, 『日本心理学会第63回大会発表論文集』, p. 429.
- Woodworth, R. S. & Schlosberg, H. 1954. *Experimental Psychology* (Revised edition). New York: Holt, Rinehart and Winston.