



# モンテカルロシミュレーションに基づく 古典的心理物理学的測定法の推定効率比較

Comparison of estimation efficiency of classical psychophysical methods  
using Monte Carlo simulation

竹井達哉<sup>1)</sup>, 鈴木颯<sup>1)</sup>, 藤原正浩<sup>1)</sup>, 牧野泰才<sup>1)</sup>, 篠田裕之<sup>1)</sup>

Tatsuya TAKEI, Shun SUZUKI, Masahiro FUJIWARA,

Yasutoshi MAKINO, and Hiroyuki SHINODA

1) 東京大学 新領域創成科学研究科 (〒 277-8561 千葉県柏市柏の葉 5-1-5, takei@hapis.k.u-tokyo.ac.jp)

**概要:** 本研究ではモンテカルロシミュレーションを用いて古典的心理物理学的測定法である恒常法, 調整法, 極限法の推定効率を定量的に比較する. 得られた結果並びに考察は, あらゆる刺激の絶対閾または弁別閾の推定における測定法並びに実験設定選択に有用な一つの定量的評価基準となる. 本研究では被験者の反応を真の心理物理学的パラメータ PSE と IU から定義される関数を用い, 乱数によって被験者の回答を模擬することで, 各古典的測定法を用いた被験者実験のシミュレーションを行う. シミュレーションは各古典的測定法について複数の実験設定別に行い, 各測定法において同一の推定精度を達成するために必要な測定回数を定量評価する.

**キーワード:** 心理物理実験, 古典的心理物理学的測定法, モンテカルロシミュレーション

## 1. はじめに

外的な刺激とそれによって引き起こされる人間の内的な感覚との間に成り立つ関係性を理解する上で, 刺激の絶対閾及び弁別閾の測定は重要なテーマの一つである. これらの閾値を測定する方法として, 現在に至るまで様々な心理物理学的測定法が提案されており, それらは古典的心理物理学的測定法と適応的心理物理学的測定法に大別される. 適応的測定法は被験者の反応に応じて閾値付近に提示刺激を随時配置する手法であるが, 古典的測定法はそうではないため, 比較的推定効率は低い. 一方で, 古典的測定法は被験者の思考が閾値の推定に影響しづらいという利点があり, 未だに広く採用される.

古典的測定法は恒常法 (Method of Constant Stimuli), 調整法 (Method of Adjustment), 極限法 (Method of Limit) から構成される. 恒常法は予め用意した刺激セットからランダムに刺激を提示する手法であり, そのランダム性から, 閾値を比較的高精度で測定できる手法だとされている. 一方で, 多くの測定を要するため測定時間が長くなり, またその結果被験者の疲労が蓄積されることによって精度が低下するという問題が生じる [1]. 調整法は被験者が自身で刺激を調整する簡単な測定手法であり, また測定回数も恒常法に比べて少なく済むため, 測定時間も比較的短い. しかし被験者の提示刺激への期待と慣れによる誤差の影響により, 精度が低く, 厳密な閾値の推定には不向きであると言われている [1]. 極限法は閾値のはるか上もしくは下から刺激

提示を開始し, 閾値方向に刺激の強さを少しずつ変化させる手法であり, 恒常法よりも短い測定時間で閾値を推定できるとされている. 一方で, 提示刺激に連続性があるため, 被験者の期待と慣れによる誤差の影響によって, 恒常法と比較して精度が低いと言われている [1].

以上のように各古典的測定法への定性的な評価は存在する. また定量的評価については, 特定の閾値測定に限り, 実験に基づいた考察が行われている (例えば, [2]). 一方で, 心理物理学的パラメータは厳密な比較が難しく, 一般的な閾値測定に対しては, 各古典的測定法同士の定量的な評価は行われていない.

そこで本研究は, 各古典的測定法において, 同一の推定精度 (もしくは, 実用上問題とならない程度の誤差) を達成するために必要な測定回数を定量評価する. 定量評価を行う上で, 本研究では被験者の反応を過去の提示刺激や回答から独立したものと仮定し, モンテカルロシミュレーションによって模擬する. そして, 各古典的測定法において主観的等価点 (Point of Subjective Equality:PSE) と不確定帯 (Interval Uncertainty:IU) がそれぞれ等価であるという仮定を置き, 各測定法の推定効率を調べ, 比較及び考察を行う.

## 2. 手法

本研究では, 予め設定した PSE と IU に基づく関数を用い, 乱数によって被験者の回答を模擬する. そして, 得ら

れた被験者の回答によって PSE と IU を推定し、その分布と分散を比較することで推定効率について考察を行う。本研究ではまず各測定法の実験設定別推定精度を比較し、次に推定精度が最も高い実験設定において各測定法間の推定効率を比較する。

## 2.1 恒常法

### 2.1.1 閾値の推定方法

恒常法では、まず事前に用意した刺激セット（通常 5 個から 9 個 [1]）からランダムに刺激を選択して被験者に提示し、それに対する被験者の二値の回答「イエス（感覚を得る、またはより大きい刺激であると感じる）」、「ノー（感覚を得ない、またはより小さい刺激であると感じる）」を記録する。そして得られた回答を、図 1 のような心理測定関数に回帰することで、閾値を推定する。なお、本研究では被験者が「イエス」と回答する確率  $p$  を式 (1) と仮定した。

$$p = f(x) = \frac{1}{1 + \exp(ax + b)}, \quad (1)$$

ここで、 $x$  は提示した刺激の値を表す。

また、恒常法による絶対閾の測定において、PSE は絶対閾に相当し、IU は被験者が 25-75 % の確率で「イエス」と回答する刺激帯の幅に相当する。弁別閾の測定においては、PSE は標準刺激に対して被験者が等しいと感じる刺激の値に相当し、IU は弁別閾の 2 倍の値に相当する。以上の PSE、IU は式 (2)、(3) のように求めることができる。

$$\text{PSE} = f^{-1}(0.5) = -\frac{b}{a}, \quad (2)$$

$$\text{IU} = f^{-1}(0.75) - f^{-1}(0.25) = -\frac{2\ln 3}{a}. \quad (3)$$

### 2.1.2 実験手順

まず、被験者の PSE と IU をそれぞれ  $\text{PSE} = 50$ ,  $\text{IU} = 20$  と設定した。これらの値から式 (1) の真の心理測定関数における  $a$ ,  $b$  を式 (2)、(3) から決定した。

次に、被験者に対して表 1 に示す 4 種類の実験設定に従って刺激提示することを仮定し、シミュレーションを行った。提示刺激の種類は 5 種類、7 種類、9 種類とし、刺激提示回数は 203 回、または 315 回である。被験者の回答は、提示刺激の値を真の心理測定関数に代入して算出した確率  $p$  で 1, 確率  $1 - p$  で 0 と出力した。

その後、提示刺激の値と被験者の回答のデータセットを基にロジスティック回帰を行い、式 (1) のパラメータ  $a$ ,  $b$

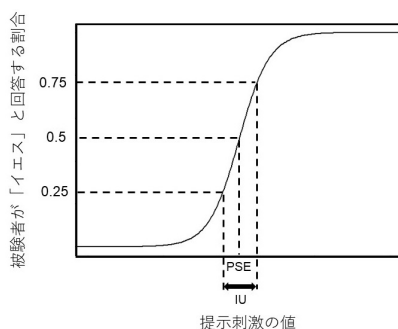


図 1: 恒常法における PSE と IU の関係

表 1: 恒常法におけるシミュレーションの詳細

パラメータ	C5_315	C7_203	C7_315	C9_315
提示刺激の種類	5	7	7	9
提示刺激の配置幅	20	13.3	13.3	10
提示刺激の下限	10	10	10	10
提示刺激の上限	90	90	90	90
各刺激の提示回数	63	45	45	35
刺激の総提示回数	315	203	315	315

を推定した。そして、推定した  $a$ ,  $b$  と式 (2)、(3) を基に PSE と IU を推定した。この一連の試行を各実験設定別に 100 回行った。

## 2.2 調整法

### 2.2.1 閾値の推定方法

調整法は、閾値と等しいと感じる刺激量为目标に被験者自身に刺激量の調整をさせ、得られた刺激量を図 2 のような正規分布関数に回帰することで、閾値を推定する方法である。

なお、調整法における PSE は絶対閾の測定では絶対閾に相当し、弁別閾の測定では標準刺激に対して被験者が等しいと感じる刺激の値に相当する。調整法の PSE は式 (4) のように表すことができる。

$$\text{PSE} = \mu(g_1), \quad (4)$$

ここで、 $g_1$  は各測定で計測された調整値の集合である。 $\mu(g_1)$  は  $g_1$  の平均を表す。

IU は一般的に調整法では想定されないが、本研究では比較のため [2] を参考に、式 (5) のように定義する。

$$\text{IU} = 2z(0.75)\sigma(g_1) = 1.349\sigma(g_1), \quad (5)$$

ここで、 $\sigma(g_1)$  は  $g_1$  の標準偏差を表し、 $z$  は  $z$  値を表す ( $z(0.75) \doteq 0.6745$ )。

### 2.2.2 実験手順

まず PSE と IU を設定し ( $\text{PSE} = 50$ ,  $\text{IU} = 20$ )、式 (4)、(5) から、被験者の調整値の確率分布である正規分布の平均と分散を決定した。次にこの正規分布によって規定される確率分布に基づく乱数によって被験者の調整値を出力した。なお、調整試行回数は 10 回 (A10)、50 回 (A50)、100 回

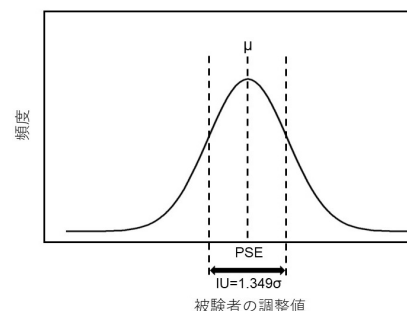


図 2: 調整法における PSE と IU の関係

(A100), 315 回 (A315) とし, それぞれで得られた調整値の平均と分散を算出し, 式 (4), (5) を基に PSE と IU を算出した. この一連の試行を各調整試行回数別に 100 回行った.

## 2.3 極限法

### 2.3.1 閾値の推定方法

極限法は, 閾値のはるか上, もしくは下から刺激の提示を開始し, 刺激提示毎に閾値方向に徐々に刺激を変化させて, 感覚を得る, もしくは失う遷移点の刺激量を記録する. そして, 得られた遷移点の刺激の値を基に閾値を推定する方法である. なお, 絶対閾と弁別閾の測定は想定する遷移状態の種類が異なる. 絶対閾の測定では提示刺激に対して「感覚がある」「感覚がない」のように 2 つの状態の遷移点に着目するが, 弁別閾の測定では 2 つの提示刺激  $X, Y$  に対し, 「 $X$  の方が  $Y$  より大きいと感じる」「 $X$  と  $Y$  が等しいと感じる」「 $Y$  の方が  $X$  より大きいと感じる」のように 3 つの状態の遷移点に着目する. 今回は被験者の回答を出力する数理モデルの都合上, 絶対閾の測定のみを行う. この場合, 絶対閾と PSE は等価である. また, IU については弁別閾の測定法でのみ想定されるパラメータであるため無視する.

なお, 極限法のシミュレーションでは恒常法と同様に被験者が「イエス」と回答する確率を式 (1) と仮定し, PSE は遷移点の刺激量を参照して式 (6) から求められる.

$$PSE = \mu(g_2), \quad (6)$$

ここで,  $g_2$  は遷移点の刺激量の集合である.  $\mu(g_2)$  は  $g_2$  の平均を表す.

### 2.3.2 実験手順

まず PSE と IU を設定し (PSE = 50, IU = 20), 恒常法と同様に式 (2), (3) から真の心理測定関数 (式 (1)) のパラメータ  $a, b$  を設定した. 次に表 2 に示す 4 種類の実験設定に従い, 刺激の下限から各刺激変化幅に応じて被験者に刺激提示することを仮定し, 被験者の回答を恒常法と同様にして, 二値で記録した. なお, 提示刺激は被験者の回答が遷移した段階で系列的な提示を止め, 刺激の上限から同様に刺激を提示した.

このように上昇系列, 下降系列に従う提示刺激を 3 種類の刺激変化幅で繰り返し, 提示刺激回数が 200 回, または 315 回を超えた後に被験者の回答が遷移した時点で刺激の提示を終了した. そして, 得られた遷移点の刺激の値から式 (6) を基に PSE を算出した. この一連の試行を各実験設

表 2: 極限法におけるシミュレーションの詳細

パラメータ	L11_200	L11_315	L21_315	L31_315
提示刺激の種類	11	11	21	31
提示刺激の変化幅	9	9	4.5	3
提示刺激の下限	5	5	5	5
提示刺激の上限	95	95	95	95
刺激の総提示回数	200 + $\alpha$	315 + $\alpha$	315 + $\alpha$	315 + $\alpha$

定別に 100 回行った.

## 3. 結果と考察

本章で行った検定は全て有意水準を 5 % とした.

### 3.1 恒常法

推定した PSE と IU の分布から得られる 95 % 確率楕円を図 3 に示す. 得られた PSE と IU の分散について, 刺激提示回数が 315 回であった C5\_315, C7\_315, C9\_315 の各組み合わせに対し  $F$  検定を行ったところ, 有意差は認められなかった.

また, C7\_203-C7\_315 間で  $F$  検定を行ったところ, PSE と IU の各分散において有意差が認められた. 図 3 並びに検定結果から, 恒常法において, 今回行った実験設定下では試行回数が多いほど推定精度が向上することがわかる.

### 3.2 調整法

恒常法と同様に, 推定した PSE と IU の分布から得られる 95 % 確率楕円を図 4 に示す. 得られた PSE と IU の分散について, 各実験設定の全ての組み合わせに対し  $F$  検定を行ったところ, A50 と A100 の推定値 IU の分散においてのみ有意差が認められなかった. 一方で, それ以外の組み合わせについては有意差が認められた. 図 4 並びに検定の結果から, 調整法において, 今回行った実験設定下では試行回数が多いほど推定精度が向上することがわかる.

### 3.3 極限法

推定した PSE の平均と標準偏差は L11\_200 が  $49.9 \pm 2.07$ , L11\_315 が  $49.9 \pm 1.66$ , L21\_315 が  $50.3 \pm 1.78$ , L31\_315 が  $49.6 \pm 1.99$  となった. L11\_315, L21\_315, L31\_315 の結果を比較すると, 提示刺激の変化幅が小さくなるほど, 推定した PSE の分散が大きくなる傾向が見られる. この傾向は, 今回の実験設定下において, 提示刺激の変化幅が小さいと閾値から遠い刺激量の提示回数が相対的に増え, 結果として閾値から遠い刺激量で遷移が発生してしまうことや, 遷移点の数が少なくなってしまうことなどが要因であると考えられる. 一方で, 推定した PSE の分散について, L11\_315, L21\_315, L31\_315 の各組み合わせに対し  $F$  検定を行ったところ, 全ての組み合わせにおいて有意差は認められなかった.

また, L11\_200-L11\_315 間で  $F$  検定を行ったところ, PSE

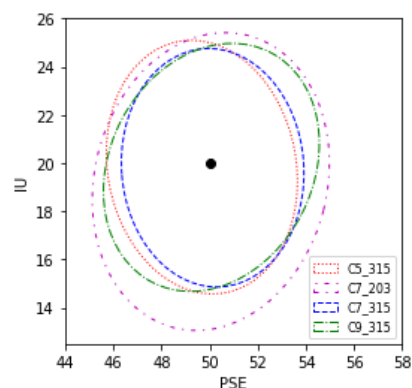


図 3: 恒常法における推定値の分布

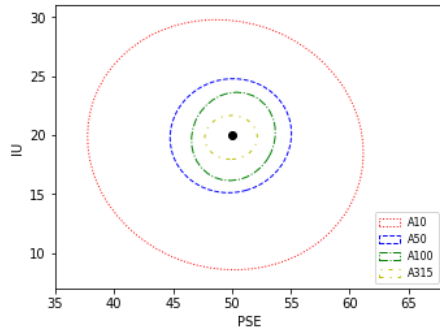


図 4: 調整法における推定値の分布

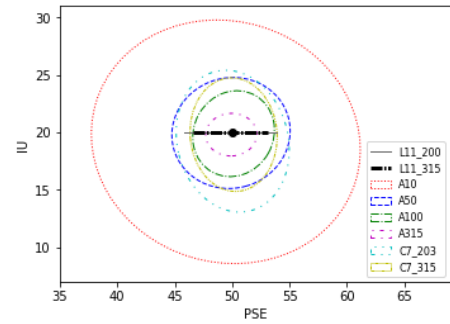


図 5: 各古典的測定法における推定値分布の比較

の分散において有意差が認められた。検定結果より、恒常法において、今回行った実験設定下では試行回数が多いほど推定精度が向上することがわかる。

### 3.4 各測定法の比較

図 5 に恒常法、極限法で最も推定値の分散が小さかった C7\_315 と L11\_315, それに対して刺激提示回数のみを変更した C7\_203 と L11\_200, また調整法の全ての場合について、それぞれ推定した PSE と IU の分布から得られる 95 % 確率楕円と 95 % 信頼区間を示している。

#### 3.4.1 恒常法と調整法の比較

推定した PSE と IU の分散について、C7\_203, C7\_315 と A10, A50, A100, A315 の各組み合わせで  $F$  検定を行ったところ、C7\_203-A50 間, C7\_315-A100 間の PSE の分散および C7\_315-A50 間の IU の分散に有意差は認められなかった。また、それ以外の組み合わせにおいては有意差が認められた。図 5 並びに検定結果より、恒常法が刺激提示回数 203-315 回の間に達成した推定精度を調整法では試行回数約 50-100 回で達成できると考察できる。

#### 3.4.2 極限法と調整法の比較

推定した PSE の分散について、L11\_200, L11\_315 と A10, A50, A100, A315 の各組み合わせで  $F$  検定を行ったところ、L11\_200-A50 間, L11\_315-A50 間および L11\_315-A100 間に有意差が認められなかった。また、それ以外の組み合わせにおいては有意差が認められた。図 5 並びに検定結果より、極限法が刺激提示回数約 200-315 回の間に達成した推定精度を調整法では試行回数約 50-100 回で達成できると考察できる。

#### 3.4.3 恒常法と極限法の比較

推定した PSE の分散について、C7\_203, C7\_315 と L11\_200, L11\_315 の各組み合わせで  $F$  検定を行ったところ、C7\_315-L11\_200 間で有意差が認められた。また、それ以外の組み合わせにおいては有意差は認められなかった。図 5 並びに検定結果より、極限法が刺激提示回数約 200-315 回の間に達成した推定精度を恒常法では約 200-315 回で十分に達成できると考察できる。

## 4. まとめ

本稿ではモンテカルロシミュレーションを用いて恒常法、調整法、極限法の推定精度を実験設定別に調べ、各古典的測定法の推定効率を比較、考察した。シミュレーションの結果、今回の実験設定下において、調整法は恒常法と極限法の刺激提示回数のおよそ 1/2 以下の試行回数で同等の推定精度を達成できることがわかった。

しかし心理物理学的測定法の推定効率を比較するにあたり、期待と慣れの誤差などのヒューマンファクターによる影響や、実験設定を決める上で必要な事前実験、測定回数あたりに要する時間など考慮すべき点が多く残る。そのため、今後は適応的測定法も比較対象として加えた上で、強制選択タスクにおける被験者の過去の回答の影響をモデル化した先行研究 [3] などを参考に、まずはヒューマンファクターや事前実験まで含めた実験モデルを各測定法別に構築し、試行回数あたりの時間の情報から定量的に測定法を選択することのできる評価基準を作成したい。

謝辞 本研究は JST CREST JPMJCR18A2 の助成を受けたものである。

## 参考文献

- [1] G. Gescheider (1997). *Psychophysics : The Fundamental Third Edition*. Lawrence Erlbaum Associates, Inc. (ゲシャイダー G. 倉片憲治・金子利佳・芝崎朱美 (訳) 宮岡徹 (監訳) (2002). *心理物理学—方法・理論・応用 (上巻) —*)
- [2] W.Chen, N. Landeveze, W. Hu, S. Ou, and P. Bourdot "Comparison Between the Methods of Adjustment and Constant Stimuli for the Estimation of Redirection Detection Thresholds", *EuroVR. Virtual Reality and Augmented Reality*, pp. 226–245, 2019.
- [3] I. Frund, F. A. Wichmann, and J. H. Macke "Quantifying the effect of intertrial dependence on perceptual decisions", *Journal of Vision*, Vol. 14, Issue 7, 9, 1–16, 2014.