

A study on the experimental efficiency of threshold determination in psychoacoustics

寺岡, 章人

<https://doi.org/10.11501/3144965>

出版情報：九州芸術工科大学，1998，博士（芸術工学），課程博士
バージョン：
権利関係：



氏名・本籍(国籍) 寺岡章人 (兵庫県)
学位の種類 博士(芸術工学)
学位記番号 甲第 24 号
学位授与の日付 平成 10 年 9 月 30 日
学位授与の要件 学位規則第 4 条第 2 項該当
学位論文題目 音響心理実験における閾値測定の効率に関する研究
審査委員会 幹事 教授 津村尚志
委員 教授 岩宮眞一郎
委員 助教授 山下由己男

論文内容の要旨

心理物理学において、心理物理定数として、閾値、弁別閾、等価値などがある。このような測定値は、感覚や知覚の特性と構造を知るうえで重要であることはよく知られている。本論文は、閾値測定の効率化について、検討を行ったのものである。閾値測定をする場合、古典的な心理測定法である調整法、極限法や恒常法から、短時間で測定を行う目的で開発された適応法がある。適応法は、実験者側の視点での測定用の効率、つまり、シュミレーションの結果で効率（真値からのずれの平均値及びそのばらつきと試行数で評価し、それらの値が共に小さいほど効率が良いと定義する）を判断するという発想で開発された。効率の良い測定法を用いて測定を行うことは重要であるが、被験者側からの視点での測定のやり易さ、つまり、被験者の負担も考慮すべき問題である。その理由として、最近の測定法の研究の流れとして、最尤推定法を用いる適応法が主流となっている。しかしこの方法は主に被験者の閾値付近のみを集中的に測定を行うため、被験者は相当な集中力を必要とし、作業が非常に困難であることが予想できる。もし、測定の途中で集中力がなくなると、データが安定しなくなる可能性もあり、逆に実験では効率が悪くなることも考えられる。

これらのことを考慮に入れ、まず、従来からあるいくつかの測定法で効率及び被験者の負担などの観点から検討を行った。その結果、他の測定法に比べ、効率が良く、かつ、被験者の負担が少ないという測定法はなかった。

本研究の目的は、被験者の負担が少なく、かつ、効率の良い測定法を開発することである。そこで、新しく RASS 法 (Rapid Adjustment of Step Size Method) を提案し、その理論及び特徴を述べ、最後にシミュレーションと聴取実験で被験者の負担を含めた総合的な効率を検討した。

第 1 章では、本研究の背景及び目的を述べた。

第 2 章では、従来からある恒常法と適応法 (PEST 法、Best PEST 法) の効率を振幅変調音の変調検出閾値実験で比較した。聴取訓練を受けた被験者群では、測定法の違いによる閾値の差はなく、実験効率も同じであることが分かった。また、被験者毎に好みの測定法が違ってくることも分かった。聴取訓練を受けていない被験者群でも、測定法の違いによる閾値

の差はなく、実験効率は同じであることが分かった。しかし、恒常法は適応法と比較すると、被験者毎の閾値がセッション間で大きく変動する被験者が多く、効率が悪いことが分かった。また、PEST 法は、測定時間が他の測定法よりも長くなり、効率が悪く、Best PEST 法の効率がよいことが分かった。一方、内観報告では、PEST 法の被験者が他の測定法の被験者より、疲労感を少なく感じていることが分かった。

第 3 章では、Best PEST 法より効率が良く、PEST 法より被験者の負担が少ない RASS 法 (Rapid Adjustment of Step Size Method) を新たに開発し、その理論について述べた。RASS 法は、刺激レベルの変化方法のアルゴリズムが非常に単純であり、測定の処理時間が速い。刺激レベルの変化がチャンスレベル付近では、起こりやすく、比較的分かりやすい刺激レベルが出やすくなっている。また、心理測定関数をロジスティック関数と仮定し、その傾きを測定値を使用することで、測定条件のパラメータが簡単に決定できるという利点がある。

第 4 章では、モンテ・カルロ・シミュレーションで適応法 (PEST 法、Best PEST 法、RASS 法) の効率を調べた。刺激レンジが狭い場合には、Best PEST 法の効率が良いが、刺激レンジが広くなるにつれ、RASS 法の効率が良くなることが分かった。このことは、実験を行う上で、閾値上の刺激レベルなら、どの刺激レベルから測定を開始しても良いという利点がある。PEST 法は、測定に要する平均試行数の標準偏差が大きな値を取るが、RASS 法は PEST 法の約半以下の標準偏差値を取る。従って、RASS 法は、測定時間が予測しやすく、実験計画を立てやすいことが分かった。そして、robust な条件の時に RASS 法が、最も効率が良いことが分かった。

第 5 章では、第 2 章と同じ変調検知閾実験を行い、3 種類の適応法 (PEST 法、Best PEST 法、RASS 法) で効率を比較した。聴取訓練を受けた被験者群では、測定法の違いによる閾値の差はなく、実験効率も同じであることが分かった。総合的な判断をすると、比較的实验計画も立てやすく、内観報告からも悪い評価がなかった RASS 法が最適であることが分かった。聴取訓練を受けていない被験者群では、測定法毎の被験者個人の閾値にばらつきがあったため、等分散の検定で閾値に有意な差があったが、被験者個人の閾値推移の分散は同じであることが分かった。内観報告の結果から、最も評価が良かったのは RASS 法である。そのため、総合的な評価をすると RASS 法が聴取訓練を受けていない被験者にも適していることが分かった。

本論文で提案した RASS 法は、他の適応法と比較して、シミュレーションで効率が良く、また、聴取実験の結果からも、被験者の負担が少なく、閾値測定に最も適していることが分かった。

論文審査の結果の要旨

人間の感覚や知覚の感度、精度などを表す基本的物理次元の指標として閾値、弁別閾、等価値などがあり、これらは感覚や知覚の特性と構造を知るうえで重要である。このよう

な心理物理定数を測定するために、これまで様々な手続きが考案されてきたが、一般の物理計測とは異なって、心理物理測定特有の組織的な刺激の呈示と測定値の最終評価のために、煩雑な手続きが常につきまとっている。従って、効率的（高い信頼度で速く）に測ることが指向され、近年、心理物理測定の効率化という観点から、適応的に測定する手法（適応法）が種々提案されている。しかし、これらの適応法は実験者側の視点での測定法の実験効率、つまり、主としてコンピュータ・シミュレーションの結果で効率改善（閾値の真値からのずれ、ばらつき、及び試行回数で評価）を目指したものであり、被験者に与える心理的苦痛や負担などには考慮しておらず、被験者側からの視点が欠けていた。

本論文では、単に実験効率が高いだけでなく、被験者に与える心理的負担の少ない、新しい適応法、RASS（Rapid Adjustment of Step Size Method）法を提案している。

第1章では、本研究の背景及び目的を述べている。

第2章では、古典的心理物理測定法である恒常法と、実験効率がよいといわれている既存の適応法（PEST法、Best PEST法）について、振幅変調音の変調検知閾実験を行い、これらの手法の効率を比較した。実験効率は、聴取訓練を受けた被験者群と受けていない被験者群とも、同じであることが分かった。しかし、恒常法は適応法と比較すると、刺激音の学習能力が悪く、また、PEST法は試行回数が他の測定法より長くなるため、結局、Best PEST法が一番効率が良いことが分かった。一方、被験者の内観報告では、PEST法が疲労感を少なく感じていることが分かった。

第3章では、Best PEST法より効率が良く、かつ、PEST法より被験者の負担が少ないRASS法を新たに開発し、その理論について述べた。RASS法は、刺激変化の規則が単純であり、測定条件のパラメータが簡単に決定でき、また、刺激レベルがチャンスレベル付近に落ち込んで、次の試行で比較的分かりやすい刺激レベルが出るようになっているという利点がある。

第4章では、コンピュータ・シミュレーションで適応法（PEST法、Best PEST法、RASS法）の効率を比較検討した。提示する刺激レベルの上限と下限の幅が広がるにつれ、RASS法の効率が良くなることが分かった。このことは、実験を開始する刺激レベルの選択範囲が広いということを含意し、RASS法の大きな利点の一つである。PEST法は、測定に要する平均試行数の標準偏差が大きな値を取るため、実験終了時間の予測が困難であり、実験の全体計画を立てにくいことが分かった。RASS法は、刺激レンジの幅、その中に存在する閾値の位置を様々に変えた条件の場合にも閾値測定は頑健であり、かつ、最も効率が良いことが示された。

第5章では、第2章と同じ実験を適応法（PEST法、Best PEST法、RASS法）について行い、これらの実験効率を比較した。聴取訓練を受けた被験者群では、実験効率は同じであったが、RASS法は内観報告からも悪い評価がなく、最適であることが分かった。一方、聴取訓練を受けていない被験者では、各測定方法毎に異なる被験者群を用いており、結果は被験者群で閾値に有意差が見られた。そこで、各測定法間のばらつきのみを見る

ため、閾値の差を補正して検討したところ、これらのばらつきは同じであることが分かった。また、内観報告から評価が良かったのは RASS 法であった。以上から、総合的な評価をすると RASS 法が聴取訓練を受けた被験者と受けていない被験者にも適していることが示された。

以上、本論文で提案された新しい心理測定手法は、単に実験効率のみを追求したのもではなく、被験者の心理的負担軽減をも視点に入れたものであり、また、この手法は他の心理実験における閾値測定の効率化にも適応可能であることから、この分野の研究の進展に寄与するところが大きいと考える。

よって、本論文は博士（芸術工学）の学位論文に値するものと、本委員会は認めた。

学力の確認の結果の要旨

最終試験においては、内観報告の検定法の妥当性、2 者強制選択法を採用した理由、RASS 法の欠点と普及させる上での問題点、等について質問があったが、十分な説明が得られた。また、論文内容、発表等より学力も十分と認められた。

公開発表会は、音響心理、視覚心理の学内関係研究室、また学外からの研究者の参加のもとで開催された。著者の発表に対して、RASS 法で 2 者強制選択法の際のチャンスレベル以下の応答を切り捨てた時の問題点、試行回数が偶数のときにバイアス（シミュレーションにおける真閾値からのずれ）が小さい理由、RASS 法の他の心理実験での汎用性、RASS 法の普及のさせ方、等について活発な質疑が行われた。いずれも著者から納得のいく説明がなされた。

よって、審査委員合議の結果、試験は合格と決定した。

もくじ

| | |
|--|-----------|
| 1 章 序論 | 1 |
| 1.1 はじめに | 1 |
| 1.2 本研究の目的 | 9 |
| 1.3 本論文の構成 | 10 |
| 2 章 恒常法と適応法の効率比較 | 12 |
| 2.1 はじめに | 12 |
| 2.2 振幅変調検知閾実験 | 14 |
| 2.2.1 実験条件 | 14 |
| 2.2.1.1 刺激音 | 14 |
| 2.2.1.2 共通する実験条件 | 16 |
| 2.2.2 聴取訓練を受けた被験者を用いた実験 | 16 |
| 2.2.2.1 実験条件 | 16 |
| 2.2.2.2 結果 | 17 |
| 2.2.3 聴取訓練を受けていない被験者を用いた実験 | 19 |
| 2.2.3.1 実験条件 | 19 |
| 2.2.3.2 結果 | 20 |
| 2.2.3.3 内観報告 | 27 |
| 2.3 まとめ | 29 |
| 3 章 Rapid Adjustment of Step Size Method (RASS 法) | 30 |
| 3.1 はじめに | 30 |
| 3.2 ルール | 32 |
| 3.2.1 刺激レベルの変更方法 | 32 |
| 3.2.2 刺激レベルの変化幅 | 34 |
| 3.2.3 求まる閾値の正応答率 | 34 |
| 3.2.4 測定条件 | 36 |

| | | |
|----------|--|-----------|
| 3.3 | まとめ | 40 |
| 4 | モンテ・カルロ・シミュレーションによる適応法の効率比較 | 41 |
| 4.1 | はじめに | 41 |
| 4.2 | 効率比較 | 43 |
| 4.2.1 | 共通するシミュレーション条件 | 43 |
| 4.2.2 | 刺激レンジが効率に及ぼす影響 | 44 |
| 4.2.2.1 | シミュレーション条件 | 44 |
| 4.2.2.2 | 結果 | 45 |
| 4.2.3 | 心理測定関数の傾き推定値が効率に及ぼす影響 | 54 |
| 4.2.3.1 | シミュレーション条件 | 54 |
| 4.2.3.2 | 結果 | 54 |
| 4.2.4 | Best PEST 法において、刺激レンジ内の閾値が刺激レンジの中心からずれた場合、効率に及ぼす影響について | 60 |
| 4.2.4.1 | シミュレーション条件 | 60 |
| 4.2.4.2 | 結果 | 61 |
| 4.2.5 | 刺激レンジ及び刺激レンジ内での閾値が存在する位置との関係をランダムに変更した場合に、効率に及ぼす影響 | 62 |
| 4.2.5.1 | シミュレーション条件 | 62 |
| 4.2.5.2 | 結果 | 63 |
| 4.3 | まとめ | 65 |
| 5 | 適応法の効率比較実験 | 68 |
| 5.1 | はじめに | 68 |
| 5.2 | 振幅変調検知閾実験 | 69 |
| 5.2.1 | 共通する実験条件 | 69 |
| 5.2.2 | 聴取訓練を受けた被験者を用いた実験 | 69 |
| 5.2.2.1 | 実験条件 | 69 |
| 5.2.2.2 | 結果 | 70 |
| 5.2.2.3 | 内観報告 | 73 |
| 5.2.3 | 聴取訓練を受けていない被験者を用いた実験 | 75 |
| 5.2.3.1 | 実験条件 | 75 |
| 5.2.3.2 | 結果 | 76 |
| 5.2.3.3 | 内観報告 | 81 |

| | | |
|---------|--|----|
| 5.2.4 | 聴取訓練を受けた被験者を用いた心理測定関数の傾き推定誤差実験 | 83 |
| 5.2.4.1 | 実験条件 | 83 |
| 5.2.4.2 | 結果 | 84 |
| 5.3 | まとめ | 86 |
| 6章 | まとめ | 88 |
| 付録 A | 第3章のシミュレーション結果 | 94 |

第 1 章

序論

1.1 はじめに

精神物理学の世界では、心理学的連続体とそれに並行した物理的連続体が仮定される。ここで、物理的連続体は音の強さ、線の長さなどの物理的な単位を用いて測定を行うことが可能であり、刺激連続体に属する。これに対し、心理学的連続体は、音の大きさ、知覚された線の長さなどであり、反応連続体に属する [1]。この反応連続体は絶えず同じ物理量で反応するとは限らない。例えば、被験者への教示、態度、動機などでも変化し [2]、物理強度にふらつきが生じる場合や神経に変動の原因がある場合もある [3]。大抵のニューロンは、外部からの刺激がなくてもランダムなパターンで発火する。刺激が存在する場合、それに上乗せする形で発火率が増加する。また、自然発火率の低いニューロンでさえ、一定の刺激に対する反応は試行毎に変化する。また、測定を行う日の体調なども影響すると思われる。

そこで、この心理量と物理量の対応関係を調べるために測定をする必要がある。なぜなら、測定をすることにより、様々な事象を数量的に記述することができ、客観的な判断、つまり、データの比較や統計処理による検討が可能となる。

その客観的な数字を得るために、これまでに様々な測定法が開発されてきた。そして、数字で統計処理を行なえるため、測定法の効率も同時に定義できる。実験で多数回の測定を行なうと、その結果、求まった閾値にばらつきが生じる。このばらつきが測定法毎に違った傾向が得られるなら、その中でも、ばらつきが小さい測定法を選ぶべきである。つまり、ばらつき (標準偏差) が効率評価をするパラメータの一つになる。また、シミュレーションで測定法を評価する場合、真値 (閾値) が分かっている。そのため、多数回のシミュレーションを行ない、その結果、求まった値の平均値と真値に差があれば、真値からのずれ (バイアス) も効率評価をするパラメータの一つになる (シミュレーションの場合のみ)。しかし、ばらつきが小さくても、試

行数 (測定に要する時間) が増加すれば、測定時間が長くなる。そのため、少ない試行数で、バイアス (シミュレーションの場合のみ)、標準偏差が共に小さな値をとる測定法が、効率が良いと定義する。

閾値推定をする場合、簡単な方法として極限法がある。極限法は主に上昇系列と下降系列の2系列を測定することで閾値を求める。上昇系列の場合、被験者は全く分からない刺激レベルから測定を行ない、始めて違いの分かる刺激が出た段階で応答をしてもらう。下降系列の場合、被験者は明らかに分かる刺激レベルから測定を行ない、始めて違いが分からない刺激が出た段階で応答をしてもらう。例えば、マスク閾値の簡単な実験で、ノイズ (レベル一定) と同時に純音を呈示し、純音の強さを徐々に小さくする。純音が聴こえなくなった時点で応答をしてもらう。そして、上昇系列及び下降系列を何度も測定し、全体の測定回数の過半数を占める刺激レベルを閾値とする場合や平均値を閾値をする場合などがある。極限法は、lapse (被験者の明らかに分かる刺激レベルでの応答ミス) に対応できないため、もし lapse があった場合、その試行が無駄になる。

古典的な測定法の中で最も洗練された測定法の一つに恒常法 (Method of Constant Stimuli) がある。恒常法は判断が分からないぐらいの刺激レベルから、明らかに分かるぐらいの刺激レベルまでの間を5から9等分程度にする。その刺激レベルを各試行でランダムに被験者に呈示し、それぞれの刺激レベルでの応答率を求める。通常、音響心理実験では、各々の刺激レベルで20試行以上の測定を行なう。そして、各々の刺激レベルでの応答率から、心理測定関数が求まる。心理測定関数には、ロジスティック関数や正規分布の累積分布関数などを使用し、閾値と心理測定関数の傾き値が得られる。

これらの古典的な測定法では、被験者の閾値付近以外の刺激レベルでも多数回の測定を行なうため、明らかに分かる刺激レベルでの試行数が無駄であるという考えが生まれた。そこで、特に閾値測定分野では、適応法 (Adaptive procedure) が開発された。適応法は、各試行で呈示される刺激レベルがそれまでの被験者の応答結果の全て、あるいは一部の履歴に基づいて逐次決められ、試行を繰り返すことによって、次に呈示する刺激レベルが閾値に次第に近づいていく測定法である。つまり、古典的な測定法は閾値以外の試行も多いが、適応法は閾値付近のみの測定を行なうため、実験時間も短時間で済み、効率が良いと言われている。

適応法は、主に以下の手続きの違いにより分類される [4]。

1. いつ刺激レベルを変えるか。
2. 次の試行レベルをどう決めるか。

3. いつ試行を終了するか。
4. 目標値 (閾値) の最終決定をどのようにするか。
5. 初期条件 (刺激レベル、刺激レベルの変化幅、心理測定関数のパラメータなど) をどう決めるか。

そして、適応法は通常 3~5 回の測定を行ない、その平均値を閾値とするのが普通である。

まず、古典的な適応法として上下法がある。これは極限法を発展させた方法である。ある試行で正応答の時、次に呈示する刺激レベルを下げ、誤応答では刺激レベルを上げる。例えば、閾値上から測定を開始すると、徐々に閾値付近まで刺激レベルが下がって行く。その後、閾値近辺の刺激レベルを上下する。そのため、極限法と比較すると、被験者の lapse にも、ある程度対応できる。しかし、1 試行毎に必ず刺激レベルが変化するため、呈示する刺激レベルが安定せず、あまり効率が良くない。そこで、さらに効率を良くするため、変形上下法 (UDTR 法)[5] が開発された。上下法は 1 試行の応答結果により、刺激レベルが変化するが、これを 2 試行以上の同じ刺激レベルの刺激に対する応答結果を用いるようにした。例えば、同じ刺激レベルで、2 回連続して正応答が出れば刺激レベルを下げる。また、誤応答が 1 回出た段階で、刺激レベルを上げるという方法である。このため、上下法では 50% の正応答閾値が算出されるが、変形上下法では、70.7% の正応答閾値が算出される。この 2 回までの応答結果で求める方法 (2 step rule) を、さらに 3 回までの応答結果 (3 step rule) を用いると、79.4% の正応答閾値が求まる (表 1.1)。

表 1.1: 主な変形上下法の刺激レベル変化パターン ('+' は正応答、 '-' は誤応答である)

| probability of positive response | down group | up group |
|----------------------------------|------------|------------------|
| 70.7% | ++ | + - or - |
| 79.4% | +++ | ++ - or + - or - |

この正応答率は、二項定理を用いて算出できる。各試行が独立であると仮定すると、正応答の出る確率が p なら、誤応答の出る確率は $(1-p)$ になる。閾値付近では正応答が出る確率と誤応答の出る確率が等しくなるため、表 1.1 より、例えば、2 step

rule の場合 (0~1 の実数解を求めると)、

$$\begin{aligned} p^2 &= p \cdot (1 - p) + (1 - p) \\ p &= 0.707 \end{aligned} \tag{1.1}$$

となる。

上下法及び変形上下法では、通常、最初の刺激レベルの変化方向が逆転するまでのステップ幅を大きくとり、その後、ステップ幅を小さくする。ステップ幅を小さくしてからの刺激レベルの変化方向が逆転した刺激レベルの平均値を閾値とする。測定の終了は、この刺激レベルの変化方向が逆転する個数をあらかじめ決めておき、その個数になった時点で終了する。この数は偶数個の場合に効率が良いといわれる。

この発展型として、1967年に M. M. Taylor と C. D. Creeman[6] によって開発された PEST 法 (Parameter Estimation by Sequental Testing) がある。大きく分けると以下の2つのルールがある。

1. 刺激レベルの変化時期について

刺激レベルの変化時期は、以下のルールによる。

$$|P \cdot T - C| < W \tag{1.2}$$

$$P \cdot T - C \leq -W \tag{1.3}$$

$$P \cdot T - C \geq W \tag{1.4}$$

- P は求める正応答率。
- T は現在呈示している刺激レベルでの試行数。
- C はそのレベルでの正応答数。
- W は deviation limits(通常、W=1 を使用する)。

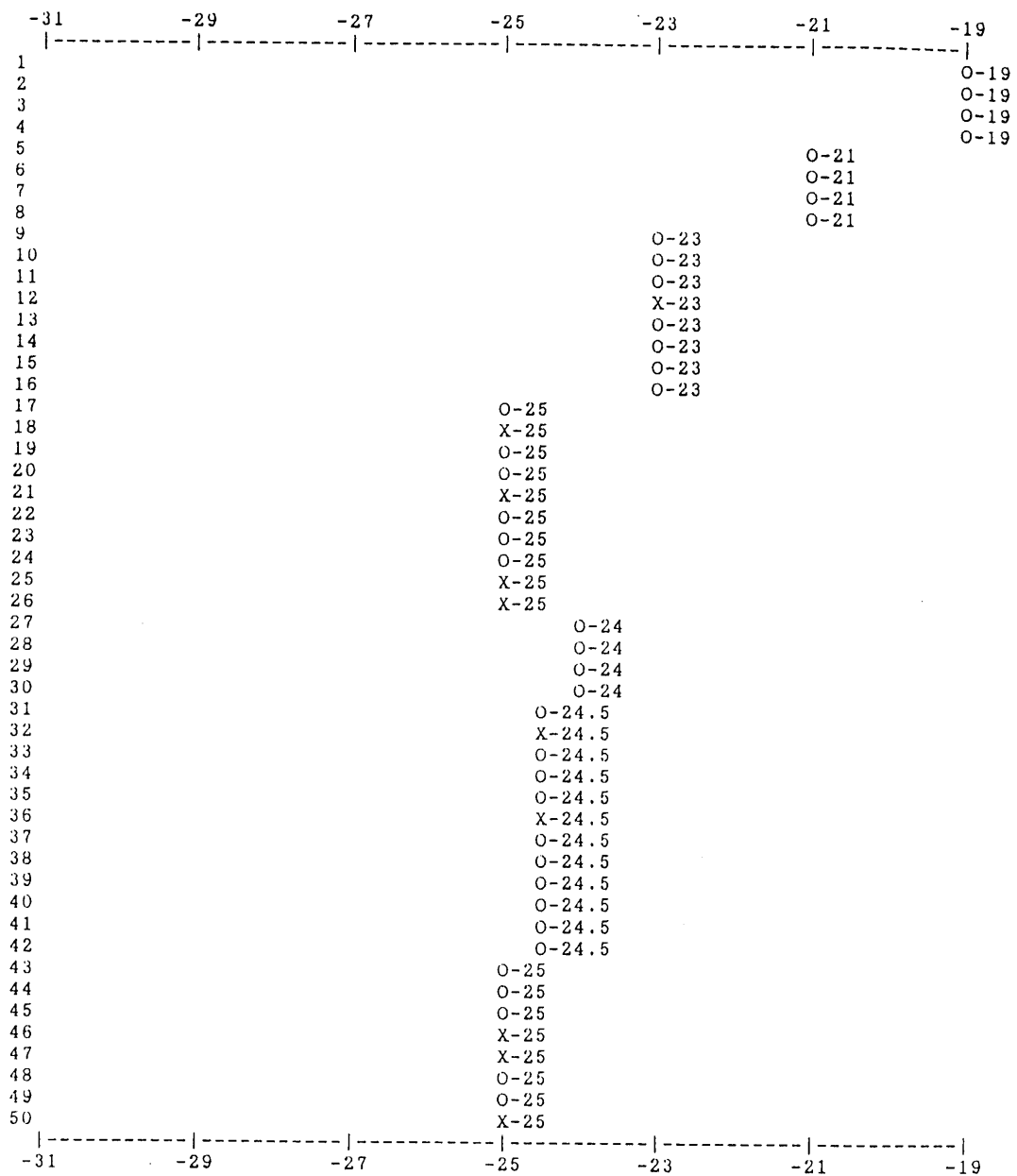
式 (1.2) なら同じ刺激レベルで、式 (1.3) なら刺激レベルを下げ、式 (1.4) なら刺激レベルを上げる。

2. 刺激レベルの変化幅について

- 次の試行ブロックで刺激レベルの増減をその前とは逆に変化させることになったとき、その刺激レベルの変化幅はその前の半分にする。この試行ブロックを刺激レベルの変化方向が逆転した1回目の試行ブロックとする。
- 刺激レベルが同じ変化方向で2回目の試行ブロックの変化幅は、その前と同じにする。

- 刺激レベルが同じ変化方向で試行ブロックが続くとき、4回目およびそれ以降のブロックの変化幅は夫々その直前の2倍にする。例えば、4回、5回と続いたときは、夫々3回目の2倍、4倍の変化幅となる。
- 刺激レベルが同じ変化方向で続く試行ブロックで3回目のブロックの変化幅は、この変化方向に逆転したときの変化幅が2倍にしたことによるものであれば変化幅は2回目のブロックと同じく、そうでなければ2回目のブロックの2倍にする。

まず、測定条件として、初期ステップサイズ、最大ステップサイズ(通常、初期ステップサイズと同じ値を用いる)、 W (deviation limit、PEST法の精度に関する値で、通常は1を用いる)、stoppingステップサイズを決定する。PEST法の終了は、ステップ幅がstoppingステップサイズになった時点で、測定が終了し、その刺激レベルが閾値になる。そのため、実験が終了するまでの試行数は一定ではない。



Threshold = -24.75 [dB]
 Trial = 50

図 1.1: PEST 法の測定例 (横軸は刺激レベル、縦軸は試行数、正応答を○、誤応答を×で示す)

この PEST 法と最尤推定法を組み合わせた Hybrid Adaptive 法 [7] が Hall によって開発された。PEST 法の刺激レベルの変化のルールに従うが、最終的な閾値は全ての試行を用いて心理測定関数を最尤推定法で当てはめて求める。しかし、この測

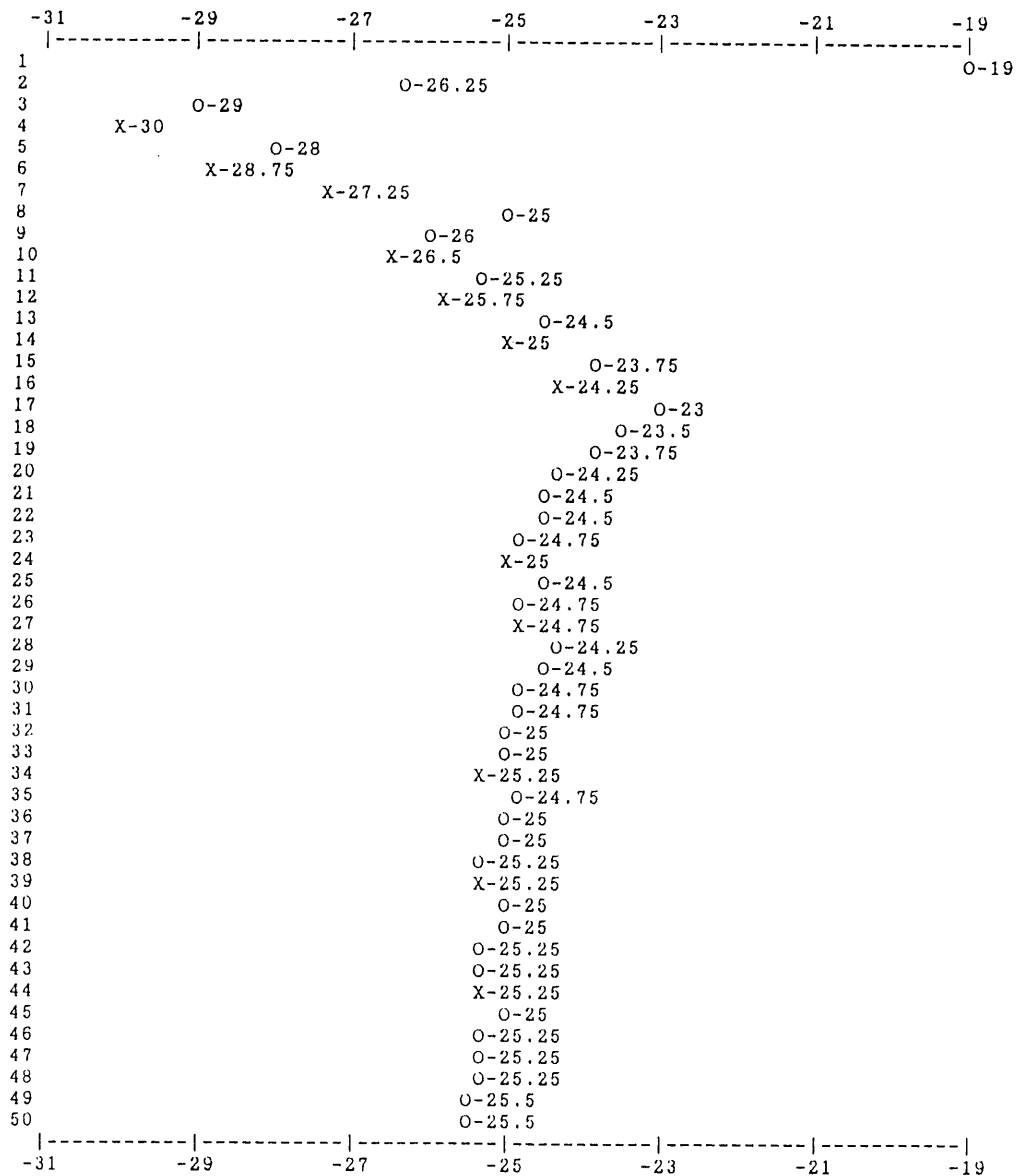
定法は著者によると 4IFC 法 (4 者強制選択法) で最も効率が良いため、時系列で刺激を呈示する聴覚心理実験では、あまり実用的でない。

この後の主な測定法の流れとして、最尤推定法を使用する測定法がメインとなっている。最尤推定法はパラメトリックな形として、心理測定関数 (ロジステック関数、Weibull 関数など) を仮定する。そして、これまでの被験者の全ての試行における応答結果を用いて、心理測定関数を当てはめる。そして、当てはめた心理測定関数から閾値を計算し、その求まった閾値を次に呈示する刺激レベルとする場合や、測定終了後に心理測定関数を当てはめて閾値を求める場合がある。

代表的なものに Best PEST 法 [8] がある。Best PEST 法は、1980 年に A. P. Pentland によって開発され、PEST 法とは全く違う手続きである。心理測定関数として式 (1.5) のロジスティック関数 (あらかじめ、傾き値を決定し、その値は一定とする) を仮定する。まず、刺激レベルの上限と下限を決定する。その間を一定のステップ幅で等分割をし、各々の刺激レベルとする。Best PEST 法は、この決められた刺激レンジ内の刺激レベルのみ呈示される。そして、被験者の応答結果の全てを用いて、毎回閾値の最尤推定をし、その求まった値 (閾値) を次に呈示する刺激レベルとする。前もって決定した試行回数になった時点で終了し、終了した時点での最尤推定値が求める閾値となる。

$$PF(x, a, b) = \frac{1}{N} + \frac{N-1}{N} \times \frac{1}{1 + \exp\{a \times (b - x)\}} \quad (1.5)$$

ただし、a: 傾き、b: 閾値、N: 強制選択肢数である。



Threshold = -25.5 [dB]

図 1.2: Best PEST 法の測定例 (横軸は刺激レベル、縦軸は試行数、正応答を○、誤応答を×で示す)

その他にも STEP 法 [9]、SIAM 法 [10] などがあるが、これらの測定法を用いた実験はほとんどないため、紹介を省く。

1.2 本研究の目的

閾値測定をする場合に、Yes/No法では被験者が絶えず「聞こえる」などの反応を示せば、測定で求まる閾値は、いくらでも低くなる。そのため、被験者の反応の偽りの影響を避けるため、Blackwell[2]によって開発された強制選択法 (forced choice method) がある。この方法は、空間的あるいは時間的に区切られた数区分の中からランダムに選ばれた一つにのみ比較刺激を呈示し、その区分がどれであるかを被験者に当てさせる。被験者は確信の有無に関わらず、強制的に一つの区分を答えなければならない。この方法では被験者が全く分からない刺激レベルで偶然に適中する率 (チャンスレベル) が一定していて、被験者の態度による反応の偽り (例えば、純音の閾値測定で、音が聴こえないのに聴こえるという反応を示すこと) に影響されない。しかし、強制選択法の選択肢の数を増やせば、チャンレベルが下がるが、1試行における時間が長くなり、被験者の負担が多くなる。また、選択肢数を増やしすぎると、音響心理実験では、同時に全刺激を呈示できないため、前回の刺激音を忘れやすく、被験者の判断が非常に難しくなる。

そこで、音響心理実験では、2IFC(two interval forced choice、2者強制選択)法が一般的に良く使用される。これは最小の選択肢数であり、測定時間が短くて済むという利点や、聴取実験では、時系列で刺激を呈示するため、選択肢数が多くなると、前回の刺激音を忘れやすく、被験者の判断が非常に難しくなるという欠点を補うためである。しかし、2IFC法ではチャンスレベルが50%と高いため、全く分からない刺激レベルでも、偶然に正応答が得られる確率が高く、効率が悪い。そこで、2IFC法を使用する場合、効率が良い測定法を用いて実験を行なう必要がある。

また、従来からある測定法は、実験者側の視点での測定法の効率、つまり、シミュレーションの結果で効率を判断するという発想であって、被験者側の負担を考慮に入れていない。これは、一般的にシミュレーションでは真値があらかじめ分かっているため、数値上での効率の差がでるが、実験では心理量を測定するため、被験者の真値が分からないので効率の差が分かりにくいからである。効率の良い測定法を用いて測定を行なうことは重要であるが、ここで、シミュレーションの結果だけを用いて、総合的な効率の判断を決定するのは問題があると思われる。その理由として、最近の測定法の研究の流れとして、最尤推定法を用いる測定法がメインとなっているが、この方法は、主に被験者の閾値付近のみを集中的に測定を行なう。ここで、被験者側からの立場になって考えると、閾値付近のみを測定する方法は、相当な集中力を必要とするため、非常に作業が困難であることが予想できる。もし、測定途中で集中力がなくなると、データが安定しなくなる可能性があり、逆に実験で

は効率が悪くなることも考えられる。そこで、被験者側からの視点での測定のやり易さ、つまり、被験者の負担も、考慮すべき問題であると思われる。被験者への負担は、測定法毎に刺激呈示のアルゴリズムが異なるため、そのアルゴリズムによって、変化する。

そこで、まず従来からあるいくつかの測定法で効率及び被験者の負担などの観点から検討を行なった。その結果、最尤推定法を使用する Best PEST 法が効率が良いことが分かった [11]。しかし、被験者の内観報告からは、PEST 法が他の測定法より、被験者の負担が少ないことが分かった。そこで、被験者側からみて、被験者への負担が少なく、かつ、効率の良い測定法を考案する必要がある。

本研究の目的は、被験者の負担が少なく、かつ、Best PEST 法と同程度、あるいは、それ以上の効率の良い測定法を開発することである。そこで、新しく RASS 法 (Rapid Adjustment of Step Size Method) を提案し、その理論及び特徴を述べ、最後にシミュレーションと聴取実験で被験者の負担を含めた総合的な効率を検討する。

1.3 本論文の構成

まず、第 2 章で、従来からある古典的な測定法である恒常法と適応法 (PEST 法、Best PEST 法) の効率比較を聴取実験で行なう。これは被験者の負担が適応法より恒常法の方が少ないであろうと考えたためである。なぜなら、恒常法はほとんど分からない刺激レベルから、明らかに分かる刺激レベルまでをランダムに呈示するため、絶えず被験者の閾値付近のみを測定しないためである。そこで、まず、聴取訓練を受けた被験者と受けていない被験者群のグループに分け、それぞれの効率を調べた。また、聴取訓練を受けていない被験者群においては、内観報告を調べることで、被験者の負担についても調べた。その結果、聴取訓練を受けていない被験者群で、閾値の安定性が悪く、恒常法の効率が悪い。また、内観からも恒常法と適応法の評価はあまり変わらなかった。そして、Best PEST 法の効率が良いことが分かった。しかし、内観からは、PEST 法が被験者の負担が最も少ないことが分かった。

以上の結果を踏まえ、第 3 章で Best PEST 法と同程度か、あるいは、それ以上の効率で被験者の負担が PEST 法より少ない適応法 (Rapid Adjustment of Step Size Method) を提案する。そして、その考え方、理論及び特徴について述べた。

第 4 章で、モンテ・カルロ・シミュレーションで、PEST 法、Best PEST 法、RASS 法の効率比較を行ない、刺激レンジが効率に及ぼす影響、心理測定関数の傾き推定値が効率に及ぼす影響、Best PEST 法において、設定した刺激レンジ内の閾値が刺激レンジの中心からずれた場合に効率に及ぼす影響、刺激レンジ及び刺激レンジ内

の閾値が存在する位置との関係をランダムに変更した場合に効率に及ぼす影響を調べた。

第5章で、聴取実験を行ない、PEST法、Best PEST法、RASS法の効率を比較する。実験効率を聴取訓練を受けた被験者と受けていない被験者群で調べる。また、被験者の内観報告も調べることで、刺激レベルの呈示方法のアルゴリズムの違いによる各測定法の被験者の負担についても比較し、被験者の負担が少ない測定法を調べた。

最後に、第6章は、本論文のまとめである。

第 2 章

恒常法と適応法の効率比較

2.1 はじめに

実験で多数回の測定を行なうと、その結果、求まった閾値にばらつきが生じる。このばらつきが測定法毎に違った傾向が得られるなら、その中でも、ばらつきが小さい測定法を選ぶべきである。つまり、ばらつき (標準偏差) が効率評価をするパラメータの一つになる。しかし、ばらつきが小さくても、試行数 (測定に要する時間) が増加すれば、測定時間が長くなる。実験での効率比較は、少ない試行数で、かつ、ばらつきの少ないほど良い。この章では、ほぼ同じ試行数で効率を比較する。従って、標準偏差が小さいほど効率が良いことになる。

被験者側の視点に立って、被験者の負担が少なく、かつ、効率が良い測定法を調べるために、まず、従来からあるいくつかの測定法で調べることにした。使用した測定法は、古典的な測定法の中で最も洗練された恒常法と適応法である。恒常法 (MCS) は、あらかじめ刺激レンジを決定し、判断が分からない刺激レベルから、明らかに分かる刺激レベルまでを通常 5 から 9 等分割し、その刺激レベルを被験者にランダムに呈示する。そのため、明らかに分かる刺激レベルでも一定の回数の試行を行なう必要があり、その刺激レベルの試行回数が無駄であるという考えが生まれた。この無駄を省き、試行数を減らすことで、実験効率を上げるという発想から、適応法が開発された。適応法は、各試行で呈示される刺激レベルが、それまでの被験者の応答結果に基づいて逐次決定し、試行を繰り返すことで閾値付近の刺激レベルを集中的に測定する。適応法で省かれた試行数は、明らかに分かる刺激レベルでの試行である。これは、実験者側からの効率を上げるという発想であり、被験者側からの測定のやり易さには繋がらないと思われる。被験者にとって、やり易い測定方法は、おおまかに分類すると以下に述べる二つの測定法の分類の内、どちらを用いるのが良いかという問題である。

- 適応法…絶えず判断が難しい刺激レベル (閾値付近) の試行を繰り返す。
- 恒常法…難しい刺激レベルから、明らかに分かる刺激レベルまでの刺激をランダムに呈示する。

そこで、被験者の負担を考慮して、あえて古典的な測定法である恒常法を比較対象とし、恒常法と適応法の効率を比較することにした。

これまでに恒常法と適応法についての効率比較を行なった論文がいくつかあり、互いに対立する結論が提出されている。シミュレーションによる効率比較は、Simpson[12] や Watson ら [13] が行なっている。Simpson はシミュレーションの結果、効率は、恒常法が良いと述べている。その反論として、Watson らは、Simpson の結果で効率が良かった理由として、「設定した刺激レンジの中心に絶えず閾値があるように条件を設定したためである。」と述べている。Watson らは、閾値を刺激レンジの中心からランダムにずらして、シミュレーションを行なった。その結果、効率は、適応法が良いと述べている。Watson らの Simpson に対する指摘は、簡単な予備実験を行なえば、被験者のおよその閾値の刺激レベルが分かり、刺激レンジの中心から閾値の設定をあえて外す必要がないため、あまり問題でないと思われる。彼らの論文の問題点は、むしろ試行数である。Simpson と Watson らは、20~100 試行で恒常法と適応法の効率比較を行なった。一般的に音響心理実験で、恒常法と適応法の実験で用いる試行数は、

- 恒常法… 100 試行以上 (各々の刺激レベルで 20 試行以上) で心理測定関数を当てはめて閾値を求める。
- 適応法… 30~80 試行を約 3~5 回の測定を行ない、その平均値を閾値とする。

である。ここで、彼らの論文で使用した試行数が、実際の音響心理実験で使用する試行数とかけ離れたものが使用されている。また、Hesse[14] は、音響心理実験を行なった結果、恒常法は PEST 法と同様な閾値のばらつきを与え、PEST 法よりも再現性が高いと述べている。

Stillman[15] は、聴取訓練を受けていない被験者群を対象として聴取実験を行なった。彼はマスク閾値の実験で、変形上下法と PEST 法で効率比較を行なった。そして、適応法の実験終了後に恒常法で測定を行ない、適応法の実験結果と比較した。その結果、恒常法で求めた閾値は、適応法と比較すると過小評価されたと述べている。彼の論文では、恒常法と適応法の効率比較をせず、適応法での試行後に恒常法と比較を行なっている。そのため、直接、恒常法と適応法を比較したわけではない。

ここでは、適応法と古典的な測定法である恒常法を用い、これまでの対立する結論も含め、聴取実験で検討する。

その内容は、

1. 各測定法の実験時間をほぼ等しくして、効率を比較する。
2. 聴取訓練を受けた被験者と受けていない被験者群に分け、測定法の違いによる閾値の差があるか、実験効率(標準偏差)は同じであるかを比較する。
3. 内観報告を取り、被験者の負担について比較する。

である。

2.2 振幅変調検知閾実験

2.2.1 実験条件

2.2.1.1 刺激音

振幅変調検知閾の測定を行ない、恒常法と適応法の効率を比較した。実験内容は、音楽のビブラート程度の早さの変調を刺激音(ノイズ)にかけ、その変調の深さ(レベル変化)が、どの程度で検知できるかを調べる実験である。刺激音の変調の有無を判断する実験であり、初めて聴取実験を受ける被験者にも、実験内容が簡単に把握できるため、この刺激音を用いて測定法の効率を比較した。

適応法は、2IFC法で同じ正応答閾値(75%)が求まる測定法を選んだ。今回使用した適応法は、適応法の中で比較的、聴取実験で良く使用されるPEST法と、コンピュータシミュレーションで効率が良いと言われているBest PEST法である。PEST法は心理測定関数を仮定せずに、使用することができ、徐々に閾値付近の刺激レベルに近付いて行く測定法である。これに対し、Best PEST法では心理測定関数をパラメトリックな形で仮定し、試行毎に最尤推定法で心理測定関数を当てはめ、その結果、求まった閾値を次に呈示する刺激レベルとする。Best PEST法は、1試行目で正応答が出ると、2試行目以降は閾値付近のみを測定する。従って、これらの適応法は、刺激レベルを変化させるアルゴリズムが全く異なる。このアルゴリズムの違いが、被験者の負担などに及ぼす影響を調べる。

刺激音は広帯域ノイズ(B & K Type 1405 Noise Generator 20~20kHz)を正弦振幅変調したgated noiseを用いた。変調周波数は8Hzの正弦波のみを用い、変調の始まりはゼロクロス点から正の位相に変化するようにした。定常部の持続時間は

500ms で刺激音の終りに 6ms のレイズドコサイン曲線で減衰を行なった。振幅の変調を変えると、刺激音の平均呈示レベルが変化する。これが、検知の手がかりにならないように、平均レベルを一定にする補正をした(式(2.1)、[16])。

$$s(t) = \frac{1}{\sqrt{1 + \frac{m^2}{2}}} \cdot [1 + m \sin(2\pi \cdot f_m \cdot t)] \cdot N(t) \quad (2.1)$$

ここで、 m は振幅変調度、 f_m は変調周波数、 $N(T)$ はキャリアノイズである。

刺激の呈示方法は AXB(3I2AFC) 法を用いた。時系列で表すと、図 2.1 のように刺激が呈示される。被験者には、真中の刺激音 (X) が前の刺激音 (A) と同じか、あるいは後ろの刺激音 (B) と同じかを答えてもらうため、標準刺激と比較刺激の区別が分かりやすい。チャンスレベル (全く分からない刺激レベルでの正答率) は 2IFC 法と同じ (50%) であるが、被験者の判断がやり易いように、AXB 法を用いることにした。

変調検知閾は、75% 正応答に対する振幅変調度 m (式(2.1) の m を $20\log(m)$ で dB 表示した) とした。また、刺激音間の空白時間は 500ms で、最後の呈示音から 2000ms 以内で被験者に応答をしてもらった。刺激音の呈示レベルは 79dB SPL で、ヘッドホン (RION AD-02) で片耳に呈示した。

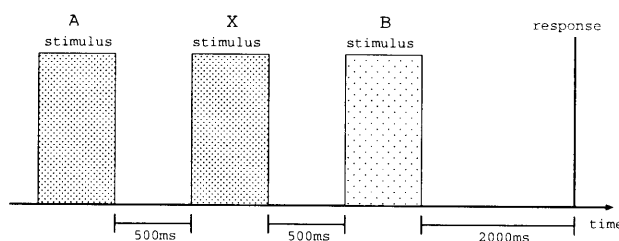


図 2.1: AXB 法の刺激呈示方法 (被験者は真中の刺激音 (X) が最初の刺激音 (A) と最後の刺激音 (B) のどちらに等しいかを答えてもらう)

各測定法の条件は、あらかじめ、聴取訓練を受けた被験者で十分な予備実験を時間をかけて行い、ほぼ等しくなるように設定した。まず、恒常法の刺激レンジ (変調の上限と下限の幅) を 12dB に決定し、PEST 法の最大ステップサイズ及び stopping ステップサイズを 50 試行前後になるように設定した。そして、PEST 法及び Best PEST 法の最小ステップサイズを同じにした (0.25dB)。Best PEST 法の刺激レンジも恒常法と同じにした。恒常法と Best PEST 法では刺激レンジの中心が予備実験より予想される閾値になるように設定し、PEST 法は予備実験の結果から求めた閾値の約 6dB 上から測定を開始した。

2.2.1.2 共通する実験条件

- 恒常法 (MCS):5 刺激レベル (3dB step)、心理測定関数にロジスティック関数 (式 1.5) を使用し、最尤推定法で閾値 b と傾き a を求めた。
- PEST 法: 初期ステップサイズ =2dB、最大ステップサイズ =2dB、 $W(\text{deviation limit})=1.0$ 、stopping ステップサイズ =0.25(dB)。
- Best PEST 法:49 刺激レベル (0.25 dB step)、心理測定関数 (ロジスティック関数) の傾き値 $a=0.6$ 。

なお、恒常法と Best PEST 法は 1 ブロック (実験で使用する最小単位の実験の区切り) を 50 試行とした。

2.2.2 聴取訓練を受けた被験者を用いた実験

2.2.2.1 実験条件

被験者は 12 人の九州芸術工科大学音響設計学科の学生 (21 から 23 歳) である。これらの学生は、本学科の正課授業として音感訓練を 2 年間にわたり、それぞれ前後学期に毎週 90 分受けている。その内容は、純音や調波複合音を用いた音の高さ、大きさ、音色、継続時間、リズムなど音の基本的性質の弁別訓練、また、音楽演奏音の一部など複雑な音を用いた音圧レベルの絶対判定及びレベル差の判定、各種フィルタによる音色変化の弁別、歪み率、残響時間の絶対判定、音楽における調性の判定、その他の訓練を含む。さらに、これらの学生は聴取実験の被験者として、これまで様々な実験に参加している [17][18][19]。

1 日に特定の測定法のみを用いると実験日の体調などの要因により、測定法毎の結果に差が出る可能性がある。これを防ぐために、1 日に全ての測定法を用いた。その方法は、1 セッション (約 1 時間) の間に、各測定法で 2 ブロックの閾値測定を行った。つまり、1 セッションで 6 ブロックの測定を行なう。そして、合計 3 セッションの測定を行った。また、全ての被験者が同じ順序で測定を行うと、実験変数 (測定法) と余剰変数 (疲労効果などの要因) が交絡する。そのため、全ての組合せ順序で実験を行い、その交絡をキャンセルするようにした。

聴取訓練を受けた被験者群での実験の目的は、各測定法で閾値に差があるかと、実験効率に差があるかを調べることである。そこで、データが安定するまで 1 日ないし 2 日の予備実験を行い、この実験の刺激音に十分慣れてもらった。その後、本試行を 3 セッション行なった。実験の始めには毎回約 30 試行の練習試行を行ってか

ら、本試行を始めた。なお、練習試行では学習効果を高めるために被験者に正誤のフィードバックを行ったが、本試行ではフィードバックを行わなかった。

2.2.2.2 結果

図 2.2 に各被験者毎の 3 日間の閾値の平均値を測定法毎に示し、図 2.3 に各測定法の被験者全員の閾値の平均値とその標準偏差を示す。ただし、恒常法は 1 日 2 ブロックの 100 試行で求めた 3 回の閾値平均値を用い、適応法は 6 回の閾値平均値を用いた。

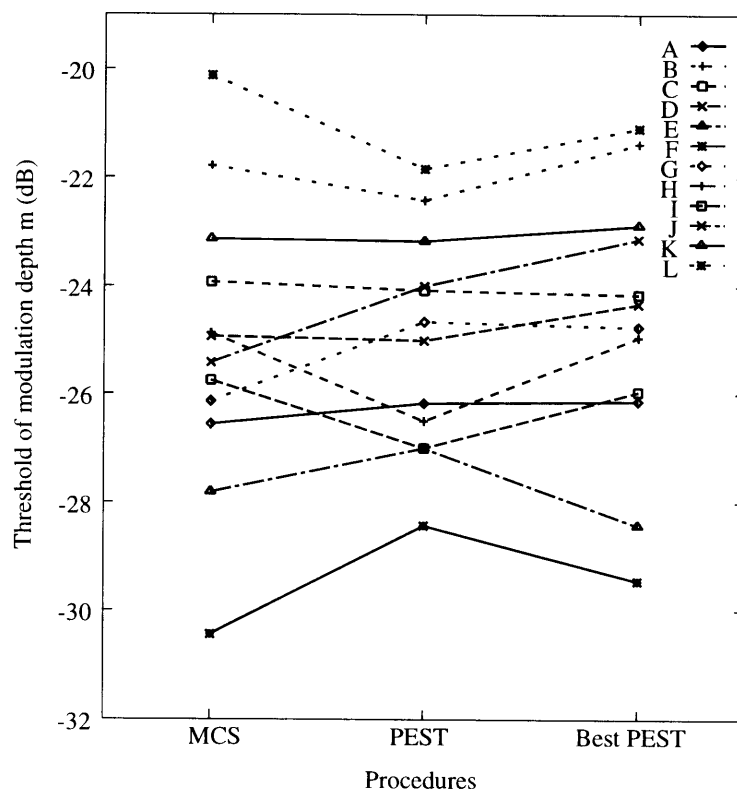


図 2.2: 聴取訓練を受けた被験者 12 人の被験者毎の変調検知閾値

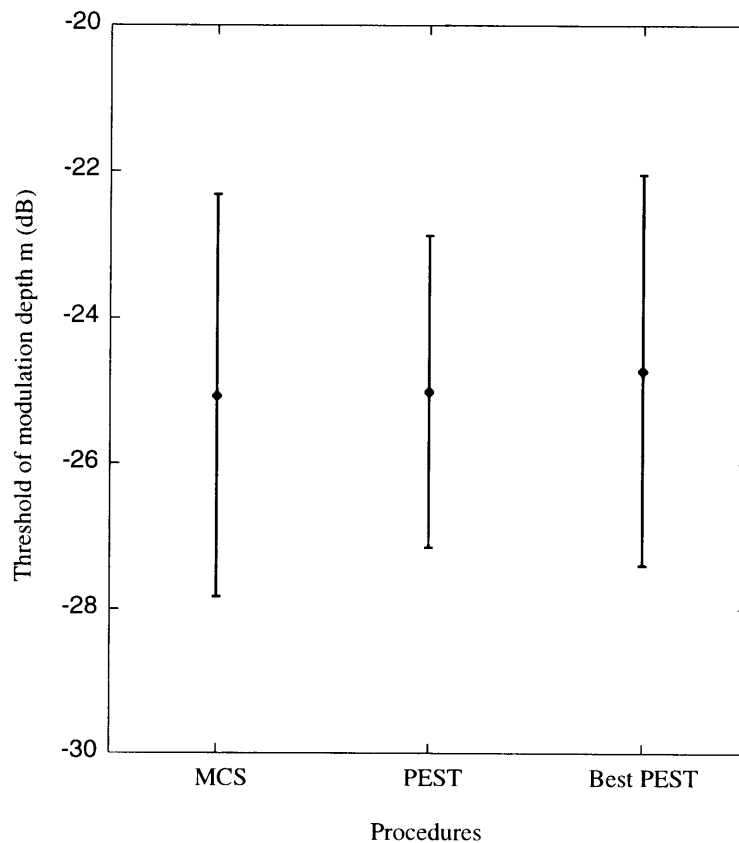


図 2.3: 聴取訓練を受けた被験者の測定法間の平均閾値及び標準偏差 (エラーバーは ± 1 標準偏差である)

恒常法の心理測定関数の傾きの平均値は、0.61(標準偏差 0.19) となり、PEST 法の平均試行数は、52.1(標準偏差 16.4) 試行であった。恒常法と Best PEST 法は、50 試行で 1 ブロックの測定が終了する。PEST 法の平均試行数(実験時間)が他の測定法とほぼ同じであるため、効率は標準偏差で評価する。そこで、Levene の等分散の検定を測定法間の閾値で行なった。その結果、標準偏差に有意な差はなかった ($F(35,144)=1.00$ 、危険率 5%)。よって、分散分析を行ない、測定法間の閾値の平均値に差があるかどうかを調べた(表 2.1)。その結果、測定法間の閾値に有意な差はなかった(危険率 5%)。聴取訓練を受けた被験者群の恒常法と適応法の実験効率は、実験時間をほぼ同じにして比較を行なったため、同じであることが分かった。

表 2.1: 聴取訓練を受けた被験者の閾値の分散分析結果

| | | | 平方和 | 自由度 | 平均平方 | F 値 | 有意確率 |
|---------|-----|-----------|----------|-----|--------|--------|-------|
| 閾値 | 主効果 | 被験者 | 832.542 | 11 | 75.686 | 58.845 | 0.00 |
| | | 測定法 | 4.427 | 2 | 2.213 | 1.721 | 0.183 |
| 2 次交互作用 | | 被験者 × 測定法 | 50.630 | 22 | 2.301 | 1.789 | 0.023 |
| | | 残差 | 185.211 | 144 | 1.286 | | |
| | | 合計 | 1125.822 | 179 | 6.290 | | |

実験終了後、被験者に内観報告をしてもらい、その検討を行なった結果、個人毎に好みの測定法があることが分かった。各測定法の内観をまとめると、

- 恒常法…実験を始めた頃は、適応法に較べるとあまり疲れないという報告があった。しかし、実験に慣れてくると、次に呈示される刺激レベルが全く予測できないので、集中しにくいという報告もあった。
- PEST 法…いつ実験が終るか分からないという不安感が大きいようであった。また、試行数が増えるにつれ不安感が増し、70~100 試行近くなると、かなり疲れるようであった。
- Best PEST 法…始めのうちは他の測定法に較べると、疲労感も大きいようであった。しかし、実験に慣れるにつれ、急激な刺激レベルの変化がないので、刺激音に集中しやすく、それほど負担もないようであった。

などである。

2.2.3 聴取訓練を受けていない被験者を用いた実験

2.2.3.1 実験条件

聴取訓練を受けた被験者群で恒常法と適応法の効率を比較した結果、実験効率は同じであることが分かった。この節では、聴取訓練を受けた被験者群と同じ実験条件を用いて、聴取訓練を受けていない被験者群の効率を比較する。この被験者群は、九州芸術工科大学の音響設計学科以外の学生で、これまでに全く聴取訓練を受けたり、聴取実験に参加したことがない。聴取訓練を受けた被験者群では、被験者全員の閾値測定に全ての測定法を使用した。しかし、聴取訓練を受けていない被験者群では、測定法の違いによる被験者毎の閾値の安定性を調べるため、1人の被験者に

は1つの測定法のみを用いて、測定を行なった。そこで、健聴者30人(19から25歳)を各測定法に無作為に10人ずつ割り振った。

実験時間は、約1時間とした(この時間区切りを1セッションとする)。1時間で練習試行(正誤のフィードバックあり)を1ブロック測定し、その後、本試行(正誤のフィードバックなし)を3ブロック測定した。また、各ブロックの間には、十分な休憩時間を取った。そして、計3日で3セッションの聴取実験を行い、各測定法で被験者の閾値の安定性などを調べた。ただし、予備試行の結果はデータに取り込まなかった。実験を始める前(1セッション目の実験開始前)に、被験者にスピーカを通して、明らかに変調していると分かる刺激音と無変調の刺激音を聞かせ、実験内容について音と図で説明を行い、実験内容を把握してもらった。また、PEST法は終了する試行数が決まってない。そのため、80試行付近において、現在呈示されている刺激レベルやステップサイズからこのまま測定を続けても、終了する見込みが少ない(刺激音の手がかりを判断する基準が分からない状態)と実験者が判断した場合、測定を途中で中断し、休憩をはさんで再度そのブロックを取り直した。

2.2.3.2 結果

セッション毎の個人データを図2.4~2.6に示す。また、それぞれの測定法で各10人のセッション間での閾値の平均値と標準偏差を図2.7に示す。ただし、恒常法は150試行ないし一部で100試行を使用し、心理測定関数を当てはめて閾値を求め、適応法は1セッションで測定した3ブロックの平均値を閾値とした。

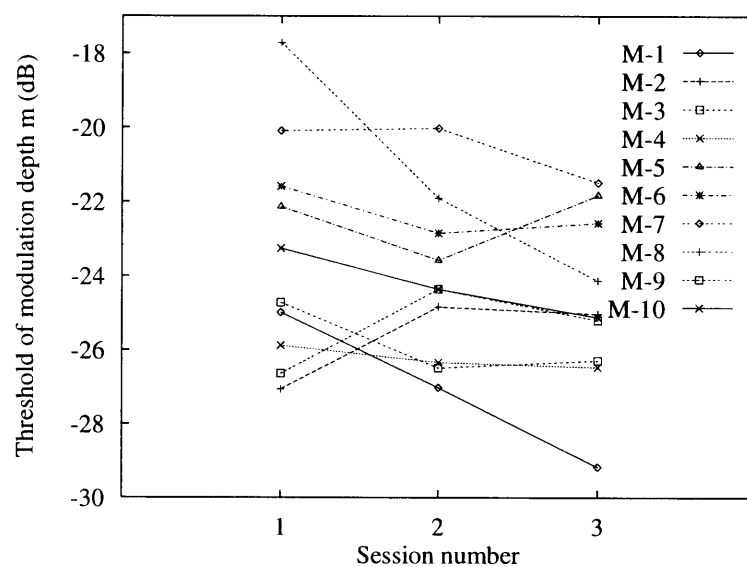


図 2.4: 聴取訓練を受けていない恒常法 (MCS) の被験者 10 人のセッション間の平均閾値

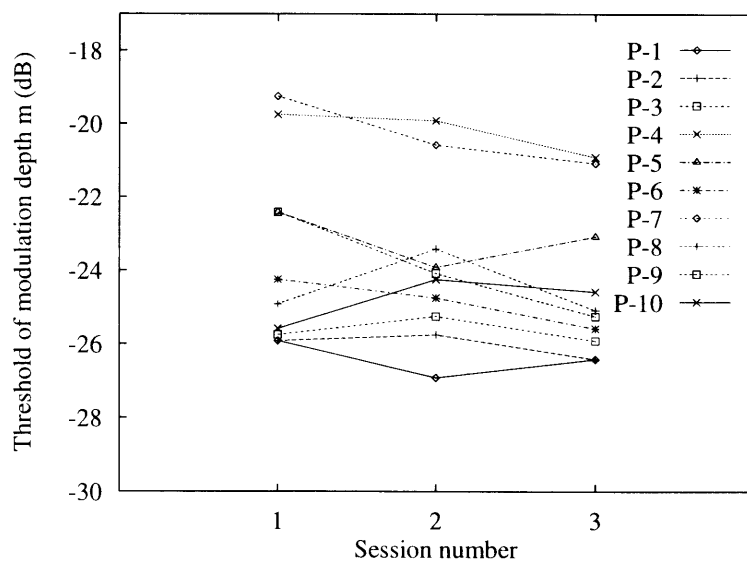


図 2.5: 聴取訓練を受けていない PEST 法の被験者 10 人のセッション間の平均閾値

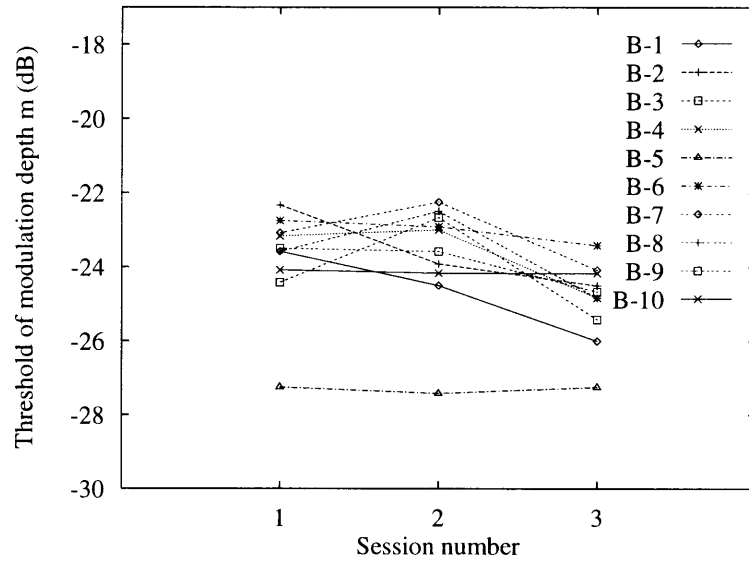


図 2.6: 聴取訓練を受けていない Best PEST 法の被験者 10 人のセッション間の平均閾値

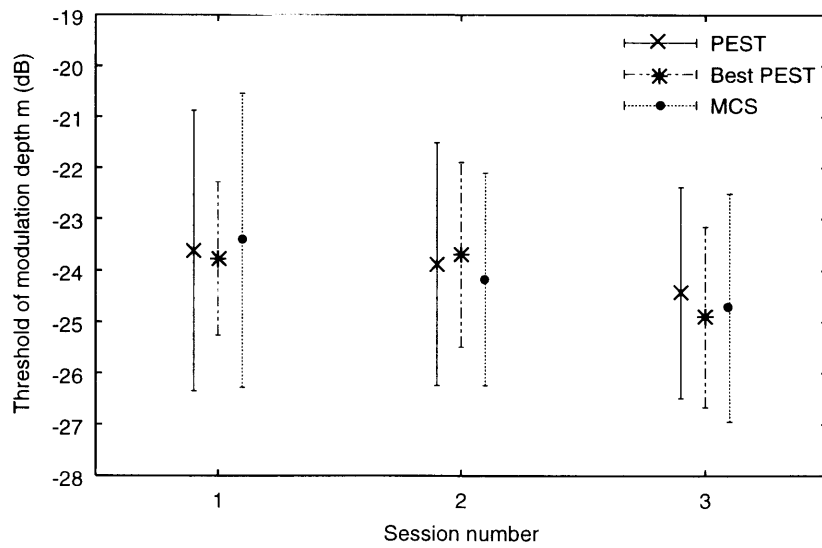


図 2.7: 聴取訓練を受けていない被験者の各測定法間のセッション間での平均閾値及び標準偏差 (エラーバーは ± 1 標準偏差を示す)

測定法の効率を調べるため、セッション間で Levene の等分散の検定を行なった。その結果、各測定法のセッション間の閾値の分散に有意差はなかった ($F(8,81)=2.04$ 、

危険率 5%)。その後、セッション間で閾値の分散分析を行った結果、閾値の平均値に有意な差 (危険率 5%) はなかった (表 2.2)。

表 2.2: 聴取訓練を受けていない被験者群の閾値に関する恒常法と適応法の分散分析の結果

| | 平方和 | 自由度 | 平均平方 | F 値 | 有意確率 |
|----------|---------|-----|-------|-------|--------|
| 閾値 グループ間 | 22.036 | 8 | 2.75 | 0.618 | 0.7602 |
| グループ内 | 360.952 | 81 | 4.456 | | |
| 合計 | 382.989 | 89 | | | |

次に、各被験者の閾値が 1 から 3 セッション間でどのように変化したのかを調べるため、被験者個人の 1 セッション目の閾値を基準 (0(dB)) として、2、3 セッション目での閾値の変化を調べた (図 2.8~2.11)。

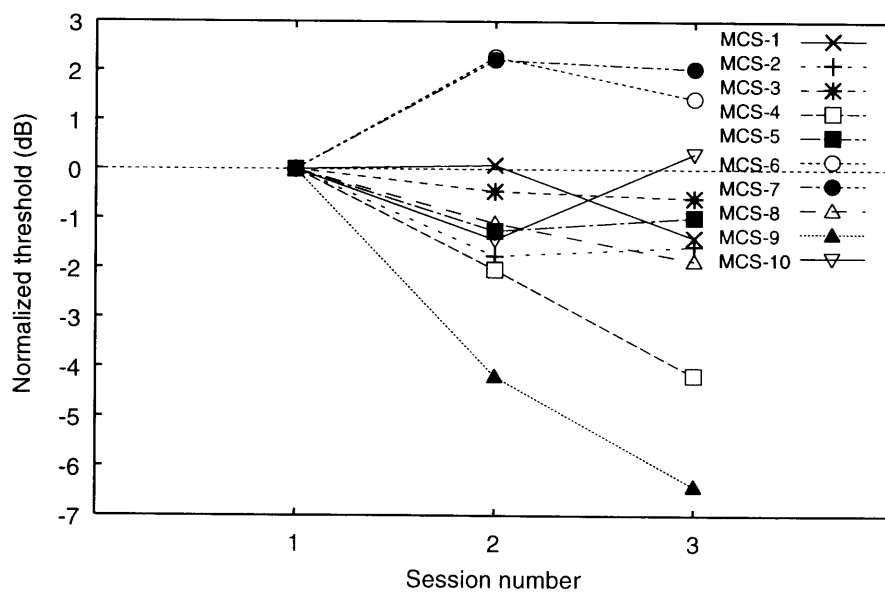


図 2.8: 1 セッション目の閾値を基準にした恒常法 (MCS) の各被験者の閾値の推移

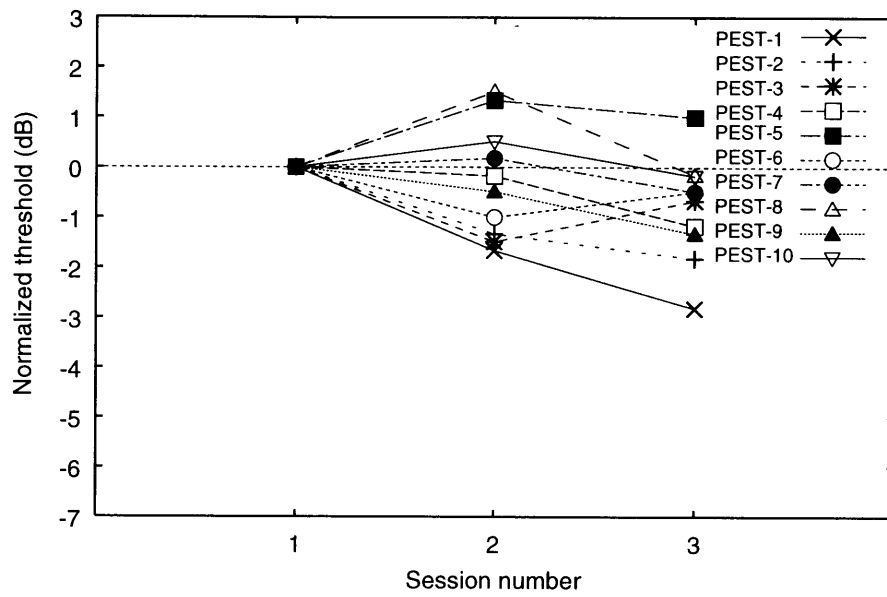


図 2.9: 1セッション目の閾値を基準にしたPEST法の各被験者の閾値の推移

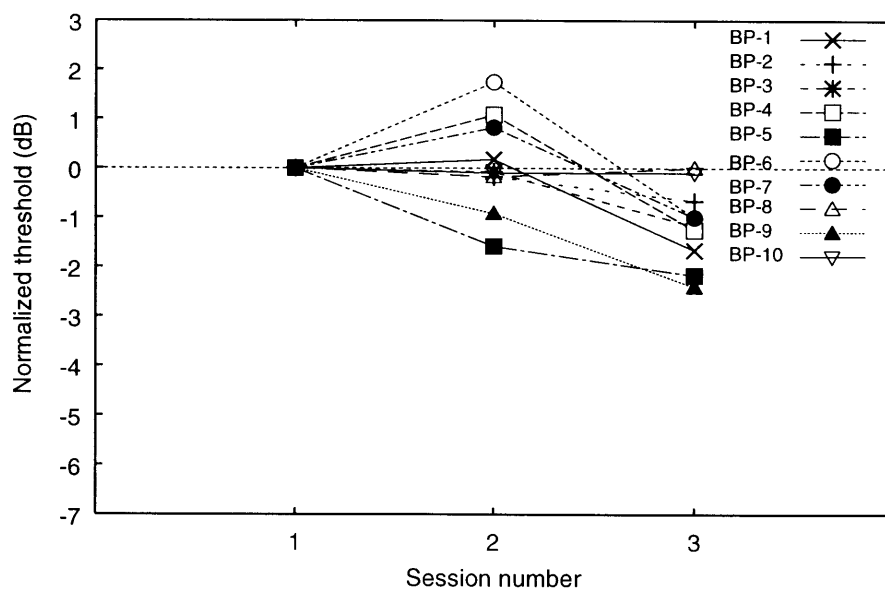


図 2.10: 1セッション目の閾値を基準にしたBest PEST法の各被験者の閾値の推移

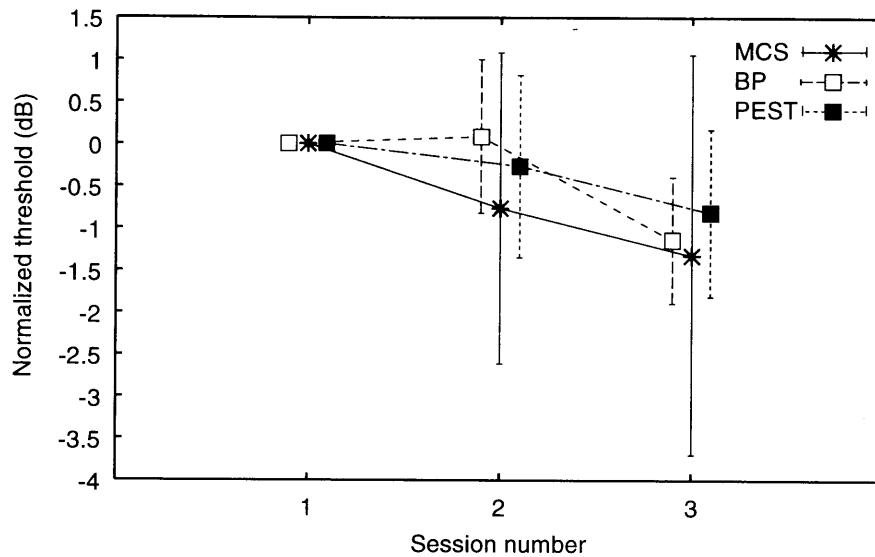


図 2.11: 1セッション目を基準にした各被験者の測定法毎の平均閾値と標準偏差 (エラーバーは ± 1 標準偏差である)

Levene の等分散の検定を被験者毎に 1セッション目の閾値で基準化した 2と3セッション目で別々に行なった。その結果、2セッション目は恒常法と適応法の閾値の分散に有意な差はなかった ($F(2,27)=1.922$ 、危険率 5%)。しかし、3セッション目は恒常法と適応法の閾値の分散に有意な差があった ($F(2,27)=3.384$ 、危険率 5%)。従って、恒常法は適応法と比較すると、被験者毎の閾値が 3セッション目で大きく変動し (図 2.8)、被験者個人の刺激音に対する学習効果が悪いことが分かった。このことは、被験者の閾値測定をする場合、適応法は最初の 1日目 (1セッション目) の測定を行なえば、被験者の閾値が分かるが、恒常法だと刺激に対する学習能力が悪いため、適応法より、測定回数が多く必要であることが分かった。よって、恒常法は適応法と比較すると、聴取訓練を受けていない被験者群を対象とした場合、効率が悪いことが分かった。

次に、2種類の適応法の効率を比較する。PEST法のセッション間での平均試行数及び標準偏差を表 2.3に示す。

表 2.3: PEST 法のセッション間での平均試行数及び標準偏差

| | Mean trial number | Standard deviation |
|-------------|-------------------|--------------------|
| 1st session | 57.4 | 12.5 |
| 2nd session | 52.5 | 12.5 |
| 3rd session | 53.8 | 11.3 |

PEST 法の平均試行数は、他の測定法 (50 試行) と比較しても、ほぼ同じである。しかし、PEST 法は測定中に被験者が刺激音の標準刺激と比較刺激の差が全く分からなくなる場合があった。例えば、80 試行前後で呈示する刺激レベルが実験開始時の刺激レベルになったり、ステップ幅が初期ステップサイズになる場合である。被験者は明らかに刺激音の手がかりを見失っており、その状態で実験を進めても、測定が終了する見込みが少ない。この場合、実験者の判断により、測定を中止した。そして、休憩をはさみ、再度そのブロックを採り直した。この実験を中断した割合が約 2 割となり、被験者は測定を中断したことを気にし、疲労した様子であった。PEST 法の問題点は、

- 実験の終了時間が予測できない。
- 被験者が刺激音の手がかりを見失うことがある。

である。そのため、PEST 法は採り直した試行数も含めると、他の測定法より実験時間が長くなったため、効率が悪いことが分かった。よって、Best PEST 法の効率が良いことが分かった。

恒常法の心理測定関数 (ロジスティック関数) の傾き値は、聴取訓練を受けた被験者群では、平均値が 0.61 (標準偏差 0.19) であり、聴取訓練を受けていない被験者群とほぼ同じ傾き値を取ることが分かった (表 2.4)。今回、Best PEST 法で心理測定関数の傾き値を聴取訓練を受けた被験者群の予備実験から求めた値の 0.6 を用いたが、本実験で求めた値とほぼ同じ値である。また、それぞれの被験者群での心理測定関数の傾き値の標準偏差も約 0.2 であるため、被験者毎に極端な傾き値を取っていない。このことは、Best PEST 法などの心理測定関数の傾き値を使用する測定法を用いる場合、本実験を始める前に実験者自身が被験者となり、心理測定関数の傾き値を求め、その求めた値を使用すれば良い。もし、実験者だけの傾き値を使用することに不安がある場合、数人の被験者に対して、予備実験を行い、その平均

値を用いれば良い。聴取訓練を受けていない被験者群の1セッション目の傾き値の平均値は、約0.6であり、2、3セッション目でも変化が余りない。よって、被験者群に関係なく、恒常法で100から150試行の測定を行なえば、簡単に心理測定関数の傾き値が求まり、その値を本実験で使用しても問題がない。

表 2.4: 恒常法のセッション間での心理測定関数の傾き値及びその標準偏差

| | Slope average value | Standard deviation |
|-------------|---------------------|--------------------|
| 1st session | 0.64 | 0.22 |
| 2nd session | 0.64 | 0.25 |
| 3rd session | 0.63 | 0.19 |

2.2.3.3 内観報告

被験者の閾値結果のみで、効率の判断をするのではなく、測定法の刺激音を呈示するアルゴリズムの違いが被験者の負担に及ぼす影響を調べるために内観報告をしてもらった。その方法は、被験者に集中力、疲労感、実験の難しさ、心理的な実験時間について、各ブロックまたはセッションの終了後に、次に述べる7段階の評価をしてもらった。

- 心理的な実験時間 (1 ブロック毎に集計)
 - 1: 非常に長い
 - 4: ちょうど良い
 - 7: 非常に短い
- 集中力 (1 ブロック毎に集計)
 - 1: 最初から集中できなかった
 - 7: 最後まで集中できた
- 疲労感 (1 ブロック毎に集計)
 - 1: 非常に疲れた
 - 7: 全然疲れなかった

- 実験の難易度 (1セッション毎に集計)

- 1: 非常に難しかった
- 4: 普通
- 7: 非常に簡単

表 2.5に各測定法の内観報告の評価値の平均値を示す。

表 2.5: 聴取訓練を受けていない被験者の内観報告の各評価の平均値

| | MCS | PEST | Best PEST |
|----------|------|------|-----------|
| 心理的な実験時間 | 3.93 | 4.00 | 4.30 |
| 集中力 | 4.52 | 5.00 | 4.81 |
| 疲労感 | 4.61 | 5.21 | 4.63 |
| 実験の難易度 | 3.83 | 3.63 | 3.50 |

今回の試行数 (1 ブロック) で、被験者は実験時間 (約 6 分) を長く感じていないため、妥当な試行数であったと思われる。また、ほとんどの人がブロックの最後まで集中力を保てないと答えている。そして、少し疲労しているようである。しかし、実験内容に関してはあまり難しくないと感じている。

ここで、各測定法で被験者の内観報告の評価値に対し、対応がない 3 条件の中央値の比較をするため、Kruskal-Wallis 検定を行なった。その結果、疲労感に有意な差があった (危険率 1%)。よって、多重比較をするために、Mann-Whitney 検定を行なった。その結果、「恒常法と Best PEST 法」では、疲労感に有意な差がなかった (危険率 5%)。そして、「PEST 法」と「恒常法と Best PEST 法」の疲労感に有意な差があった (危険率 1%)。そのため、PEST 法が最も疲労感が少ないということが分かった。しかし、被験者が刺激音の手がかりを見失う可能性も高く、実験を途中で中断し、測定を取り直した割合も多かったため、一概に聴取訓練を受けていない被験者に最も適しているとは言いがたい。

被験者の負担を考え、恒常法を比較対象に入れたが、恒常法の内観は、適応法とほぼ変わらないという結果が得られた。ここで、適応法と恒常法の刺激呈示のアルゴリズムの違い、つまり、「適応法では閾値付近を集中的に測定する」ことと、「恒常法では明らかに分かる刺激レベルから、全く分からない刺激レベルまでをランダムに呈示する」ことの違いは、聴取訓練を受けていない被験者群では、ほとんど被

験者の負担の差がないことが分かった。この理由として、適応法は被験者の正応答率が約 100% 付近から測定を始め、75% 前後の試行を繰り返す。恒常法は約 50~100% の試行をランダムに測定する。ここで、恒常法は閾値以上の刺激レベル (約 85~100%) の試行の繰り返しは、易しく感じるが、逆に閾値以下の刺激レベル (約 50~65% 付近) が非常に難しいため、内観からも良い結果が得られなかったと思われる。

2.3 まとめ

被験者の負担が少なく、かつ、効率の良い測定法を調べるために、従来からある測定法 (恒常法と適応法 (PEST 法、Best PEST 法)) で、聴取訓練を受けた被験者と受けていない被験者群に分け、変調検知閾実験で比較をした結果、以下のことが分かった。

- 聴取訓練を受けた被験者群では、測定法の違いによる閾値の差はなく、実験効率も同じであることが分かった。また、被験者毎に好みの測定法が違うことが分かった。内観報告では、恒常法は実験に慣れてくると、次に呈示する刺激レベルの予測がつかないので、集中しにくいという意見もあった。PEST 法は実験の終了時間が分からないため、被験者に不安感を与えやすい。Best PEST 法は刺激レベルの変化が余り無いため、実験に慣れると刺激音に集中しやすいという意見もあった。
- 聴取訓練を受けていない被験者群では、測定法の違いによる閾値の差はないことが分かった。しかし、恒常法は適応法と比較すると、被験者毎の閾値がセッション間で大きく変動する被験者が多く、刺激音に対する学習効果が悪いため、効率が悪い。それに対して、内観報告からは、恒常法の評価は適応法とあまり変わらない結果が得られた。また、PEST 法は他の測定法より取り直したブロックが多くなり、その結果として、測定時間が長くなり、効率が悪い。よって、Best PEST 法が実験効率が良いことが分かった。しかし、内観報告では、PEST 法の被験者は他の測定法の被験者より、疲労感を少なく感じていることが分かった。

第 3 章

Rapid Adjustment of Step Size Method (RASS 法)

3.1 はじめに

第 2 章で、恒常法と適応法の振幅変調検知閾実験を聴取訓練を受けた被験者と受けていない被験者群に分け、実験による効率の比較をした。その結果、聴取訓練を受けていない被験者群では、恒常法の閾値の安定性が悪く、恒常法の効率が悪いことが分かった。そして、聴取訓練を受けていない被験者群で、Best PEST 法が効率が良かったが、内観では、PEST 法より疲労に関する評価が悪かった。

そこで、実験効率が従来からある測定法より良く、かつ、被験者の負担も少ない測定法を新たに開発することにした。

一般的に、聴覚心理実験では 2IFC 法 (2 者強制選択法) が使用されている。その理由は、選択肢数が増えると、実時間での刺激音を呈示するため、前回の刺激音を忘れやすく、判断が難しくなるためである。そこで、2IFC 法で効率が良く、かつ、被験者の負担の少ない RASS 法 (Rapid Adjustment of Step Size Method) を開発した。

あらかじめ、その特徴を記述すると、

- 測定条件が簡単に、心理測定関数 (ロジステック関数) から決定される。
- 刺激レベルの変化方向のアルゴリズムが単純である。
- 刺激レンジを設定しない。

である。

特定の刺激レベルで多数回の測定を行なうと、その刺激レベルでの正応答率が得

られる。この時に心理測定関数のおよその傾き値が分かっているならば、心理測定関数からおよその閾値を推定することが出来る。この考え方を元に RASS 法を考案した。

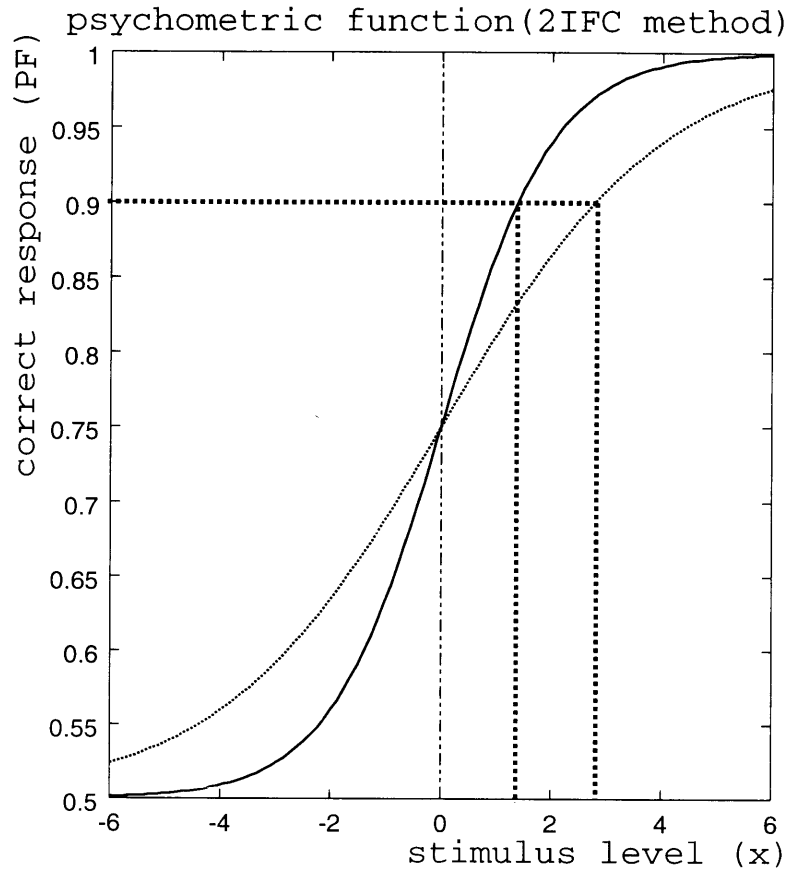


図 3.1: 2IFC 法の心理測定関数 (ロジスティック関数) の例

図 3.1では、傾き値が違う 2つの心理測定関数 (ロジスティック関数、式 (1.5)) を示している。太点線で示される刺激レベルでの正応答率が分かれば、それに心理測定関数を当てはめることにより、太点線で示される刺激レベルから閾値 (75% 正応答率) までの距離 (横軸) が算出できる。その距離を刺激レベルから引くことで閾値が求まる。

そこで、特定の刺激レベルの正応答率を調べるために、次節でその方法について述べる。

3.2 ルール

3.2.1 刺激レベルの変更方法

- 同じ刺激レベルで誤応答が2回出た場合、その刺激レベルでの試行は終了する。
- 同じ刺激レベルでの試行数は最大5回までである。

上記の刺激レベルが変更される応答パターンを図 3.2 に示す。

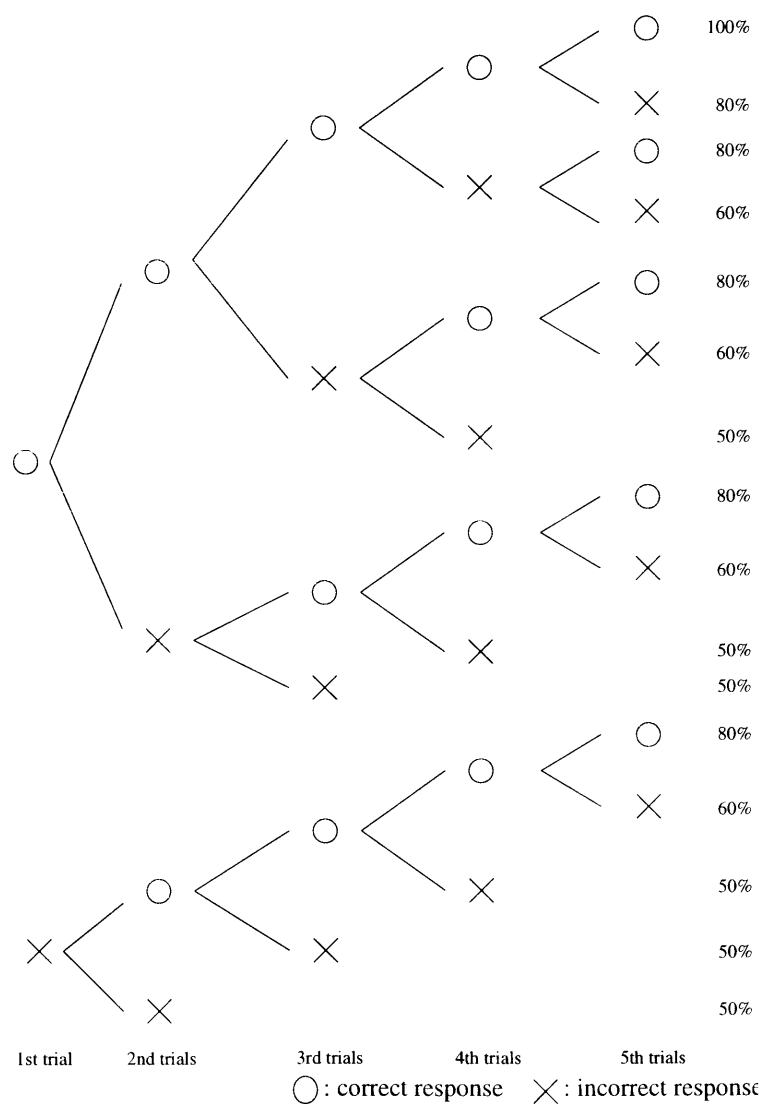


図 3.2: 刺激レベルの変化パターン (○は正応答、×は誤応答を示す)

変形上下法は、同じ刺激レベルで1回の誤応答が出た段階で、現在呈示している刺激レベルが上がる。この刺激レベルの変化アルゴリズムだと、lapse(明らかに分か

る刺激レベルでの応答ミス)に対応できない。そこで、RASS法は同じ刺激レベルの試行で、誤応答が2回出た段階で刺激レベルを上げることにした。また、同じ刺激レベルでの最大試行数を5回に限定しているため、PEST法と比較すると、刺激レベルの変化が起こりやすくなっている。

次に、同じ刺激レベルでの最大試行数を5回までにした理由を述べる。RASS法は、ある刺激レベルのおよその正応答率を求め、心理測定関数の傾き値を利用して、閾値を推測する測定法である。そこで、同じ刺激レベルで試行を繰り返し、正応答率を算出する。図3.2(6試行目は図にない)より、表3.1に同じ刺激レベルでの最大試行数を変数とした場合の求まる正応答率を示す。

表 3.1: 同じ刺激レベルでの最大試行数を変数とした場合の正応答率パターン

| 最大試行数 | 求まる正応答率 |
|-------|------------------------|
| 2 | 100%, 50% 以下 |
| 3 | 100%, 67%, 50% 以下 |
| 4 | 100%, 75%, 50% 以下 |
| 5 | 100%, 80%, 60%, 50% 以下 |
| 6 | 100%, 83%, 67%, 50% 以下 |

RASS法は、2IFC法で効率の良い測定法を開発することを目的としている。そこで、2IFC法のチャンスレベル(全く分からない刺激レベルでの正応答率)は50%であるため、正応答率が50%以下に算出される正応答率は、全て50%正応答率と仮定する。まず、求める正応答率を75%付近(心理測定関数の傾きがもっとも急な地点であり、2IFC法で50~100%の midpoint にあたる)にしたい。そして、正応答率を利用して、閾値までの刺激レベルの距離を算出する。そのため、最大試行数が4回以下の場合、求まる正応答率は3種類以下と少なく、あまり正確な正応答率が求まらない。また、75%付近を閾値とするため、最大試行数が4回の場合、75%の正応答率が出た場合、刺激レベルの変化が起こりにくい。また、最大試行数が5と6回の場合、求まる正応答率の数は両方とも同じ4つである。効率は試行数も関係し、少ない試行数ほど効率が良い。よって、最大試行数を5回にした。また、75%付近の正応答閾値を求めたいため、正応答率パターンが75%以上では次に呈示する刺激レベルを下げ、75%以下では上げることにした。

図3.2より、正応答率が50%となるパターンが多く、全体の約40%を占めている。

そのため、RASS法は判断が困難な刺激レベルでは、50%の正応答率が出やすくなっている。50%の正応答率が出た場合、正応答率から閾値(75%正応答率)までの距離を、次に呈示する刺激レベルの変化幅にするため、比較的、分かりやすい刺激レベルが次の試行で呈示されやすい。

3.2.2 刺激レベルの変化幅

それぞれの刺激レベルでの正応答率を求め、心理測定関数の傾き値を利用し、次に呈示する刺激レベルの変化幅を決定する。正応答率が50%及び100%の場合、閾値を算出する時に心理測定関数を利用すると、正応答率が99%及び51%付近では正応答率の変化幅に対して、刺激レベルの幅が非常に広く、閾値までの刺激レベル幅(距離)が求まらない。そこで、近似式を利用することで、この問題を解決する。その方法は、心理測定関数の75%点の接線を通る1次関数を利用することである。これは、心理測定関数が75%付近では、ほぼ直線と近似できることを利用した。正応答率が50及び100%での近似はおおざっぱであるが、それに対し、75%付近では一次関数でほぼ近似できる。また、一次関数で近似をすることで、各正応答率から75%正応答率の差を比で表すことが出来る。その比を各正応答率での刺激レベルの変化幅で使用するため、測定法のアルゴリズムが単純になり、RASS法の利点の一つとして挙げられる。

3.2.3 求まる閾値の正応答率

RASS法はそれぞれの正応答率での刺激レベルの変化幅を変更しているため、それぞれの刺激レベルの変化パターン(確率事象)に重み付をし、求まる閾値の正応答率を算出する。75%を通る1次関数で心理測定関数を近似しているため、それぞれの正応答率での重み付(変化幅の比)は、

$$\begin{aligned} & (100\%の正応答率が出た場合の刺激の変化幅) : \\ & (80\%の正応答率が出た場合の刺激の変化幅) : \\ & (60\%の正応答率が出た場合の刺激の変化幅) : \\ & (50\%の正応答率が出た場合の刺激の変化幅) \\ & = (75\% - 100\%) : (75\% - 80\%) : (75\% - 60\%) : (75\% - 50\%) \\ & = -25 : -5 : 15 : 25 = -5 : -1 : 3 : 5 \end{aligned} \tag{3.1}$$

となる。この重み付の比を各正答率の刺激レベルの変化パターンに当てはめる。

この重み付の比を利用する測定法にWeighted Up-Down法[20]があり、RASS法の閾値の正応答率の求め方と同じ方法であるため、紹介をする。例えば、上下法の

ステップサイズを正応答が出た場合と誤応答が出た場合に違うステップサイズを使用することで、任意の正応答率を求める方法を紹介している。その方法は、正応答が出る確率を p とすると、二項定理により、誤応答が出る確率は $(1-p)$ になる。そして、正応答率が出た場合の刺激レベルの変化幅を Sup 、誤応答が出た場合の刺激レベルの変化幅を $Sdown$ とすると、閾値付近の収束状態では、式 (3.2) の関係式が導かれる。

$$Sup \cdot p - Sdown \cdot (1 - P) = 0 \quad (3.2)$$

例えば、75% 正応答率を求める場合、 $p=0.75$ を式 (3.2) に代入すと、

$$Sup/Sdown = 1/3 \quad (3.3)$$

が求まる。つまり、

$$\begin{aligned} &(\text{正応答が出た場合の重み付}) \cdot (\text{正応答の事象が起こる確率}) + \\ &(\text{誤応答が出た場合の重み付}) \cdot (\text{誤応答の事象が起こる確率}) = 0 \end{aligned} \quad (3.4)$$

となる。Weighted Up Down 法では上下法の刺激レベルの変化ルールに従うが、正応答が出た時には、1 ステップ刺激レベルを下げ、誤応答が出た時に 3 ステップ刺激レベルを上げることで 75% の閾値を求める。

この Weighted Up-Down 法と同じ方法で、RASS 法の閾値の正応答率を求める。そこで、それぞれの刺激レベルの変化パターン (確率事象) に重み付をする。

$$\begin{aligned} &(100\% \text{の重み付}) \cdot (100\% \text{の事象が起こる確率}) + (80\% \text{の重み付}) \cdot (80\% \text{の事象が起こる確率}) \\ &+ (60\% \text{の重み付}) \cdot (60\% \text{の事象が起こる確率}) + (50\% \text{の重み付}) \cdot (50\% \text{の事象が起こる確率}) = 0 \end{aligned} \quad (3.5)$$

これを満たす 0~1 の実数解が求める正応答率となる。

そこで、今回のそれぞれの確率事象 (図 3.2) に重み付を行ない、0 ~ 1 の間の実数解を解く。

$$\begin{aligned} &-5 \cdot p^5 - 1 \cdot 5p^4(1 - p) + 3 \cdot 4p^3(1 - p)^2 + 5 \cdot \{3p^2(1 - p)^2 + 2p(1 - p)^2 + (1 - p)^2\} = 0 \\ &-12p^5 + 14p^4 + 8p^3 - 5 = 0 \\ &p = 0.75236 \end{aligned} \quad (3.6)$$

となり、75.2% の正応答閾値が求まる。また、式 (3.6) の関数式のグラフは、図 3.3 となる。

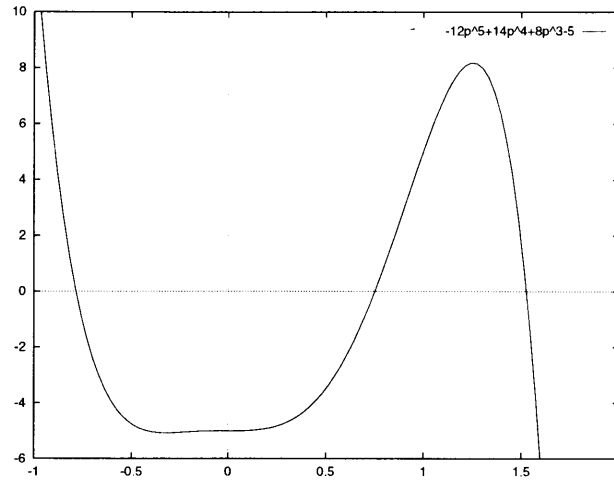


図 3.3: RASS 法の関数式

3.2.4 測定条件

以上、新しく開発した測定法の理論を述べてきたが、シミュレーションや実験を行なう時に、測定条件のパラメータを決定する必要がある。このパラメータを誤って決定すると、試行数が増加し、効率も悪くなる。そのため、測定条件のパラメータは通常、予備実験の段階で実験者が十分な時間をかけて最適なパラメータを探し、その求めた値を本実験で他の被験者にも適用する。測定法によって、このパラメータ数が増加し、パラメータ数が多くなるほど、予備実験の時間も長くなる。そこで、これらのパラメータ条件を最適に決定する方法を提案する。

RASS 法を考えた出発点は、心理測定関数の傾きと、ある刺激レベルでの正応答率から閾値 (75% 正応答率) までの刺激レベル幅を算出し、その幅を次に呈示する刺激レベルとすることである。そこで、心理測定関数にロジステック関数 (式 (3.7)) を使用し、その傾き値から、各正応答率から閾値までの刺激レベル幅を決定する方法を取った (図 3.4)。第 2 章の実験結果より、聴取訓練を受けた被験者と受けていない被験者群の心理測定関数の傾き値の標準偏差は、約 0.2 であった。つまり、心理測定関数の傾き値は、被験者が違って、あまり変化しないことを利用している。

まず、ロジステック関数の傾きを a とし、 $PF(0) = 0.75$ を通る接線を求める。

$$PF(x) = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{1 + \exp(-a \cdot x)} \quad (3.7)$$

$$\frac{dPF(x)}{dx} = \frac{a}{2 \cdot \exp(a \cdot x) \cdot \{1.0 + \exp(-a \cdot x)\}^2} \quad (3.8)$$

$$\frac{dPF(0)}{dx} = \frac{a}{8} \quad (3.9)$$

故に、 $PF(0)=0.75$ を通る接線は、

$$y = \frac{a}{8} \cdot x + 0.75 \quad (3.10)$$

となる。この一次関数とそれぞれの正応答率までの距離が、次に呈示する刺激のステップサイズとなる (図 3.4)。

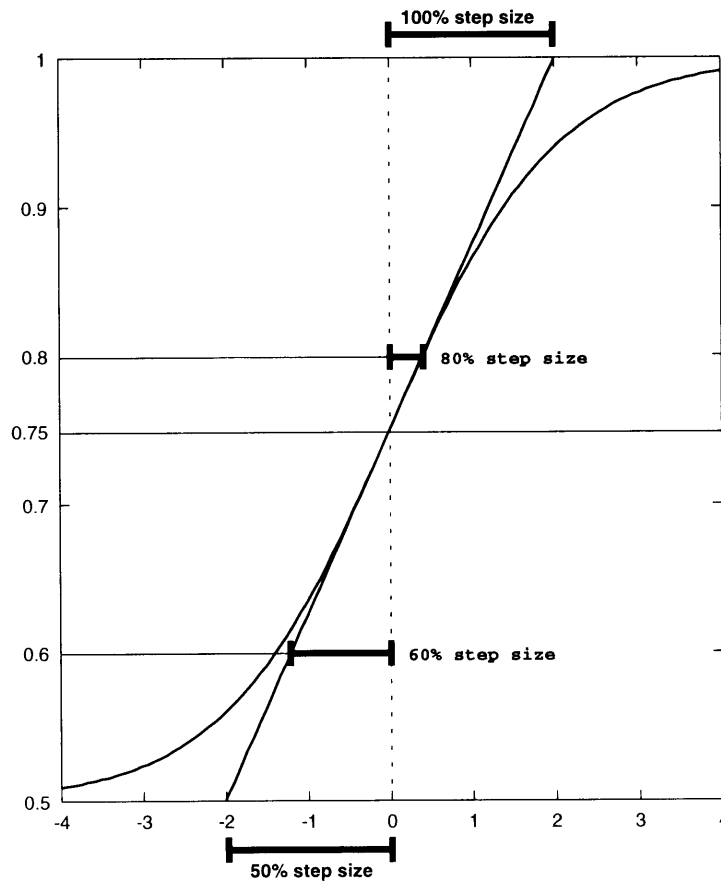


図 3.4: それぞれの正応答率での刺激レベルの変化幅

RASS 法は測定条件を決定する時に、ロジスティック関数の 75% での接線の傾き値を使用する。恒常法で求めた測定結果に、正規分布の累積分布関数を当てはめた場合でも、75% 付近では、ロジスティック関数との違いはほとんどないため、測定条件を決定する上での問題はない。図 3.5 に 2IFC 法でロジスティック関数と正規分布の累積分布関数の対応関係を示す。

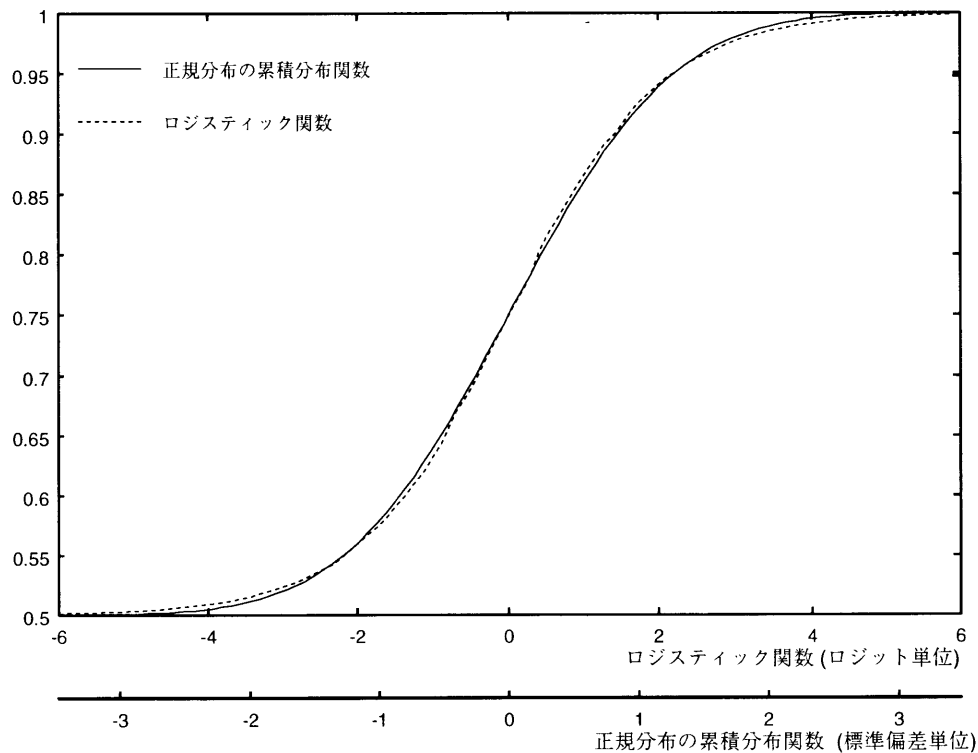


図 3.5: 2IFC 法でロジスティック関数 (傾き値 1、ロジット単位) と正規分布の累積分布関数 (標準偏差) の対応関係

RASS 法は、あらかじめ刺激レベルの変化方向が逆転する個数を決めておき、その数になった時点で、測定を終了する。そして、求める閾値は刺激レベルの変化方向が逆転した時点での刺激レベルの平均値で決定する。ただし、変形上下法では最初に刺激レベルの逆転方向が逆転した刺激レベルまでのデータを捨てることが多いが、RASS 法では捨てずに全てを使用する。

RASS 法の特徴は、測定を行う前に、あらかじめ各正応答率でのステップ幅を計算で求めることができる。そのため、複雑な計算を測定中に行う必要はない。ここに RASS 法の測定例を示す。

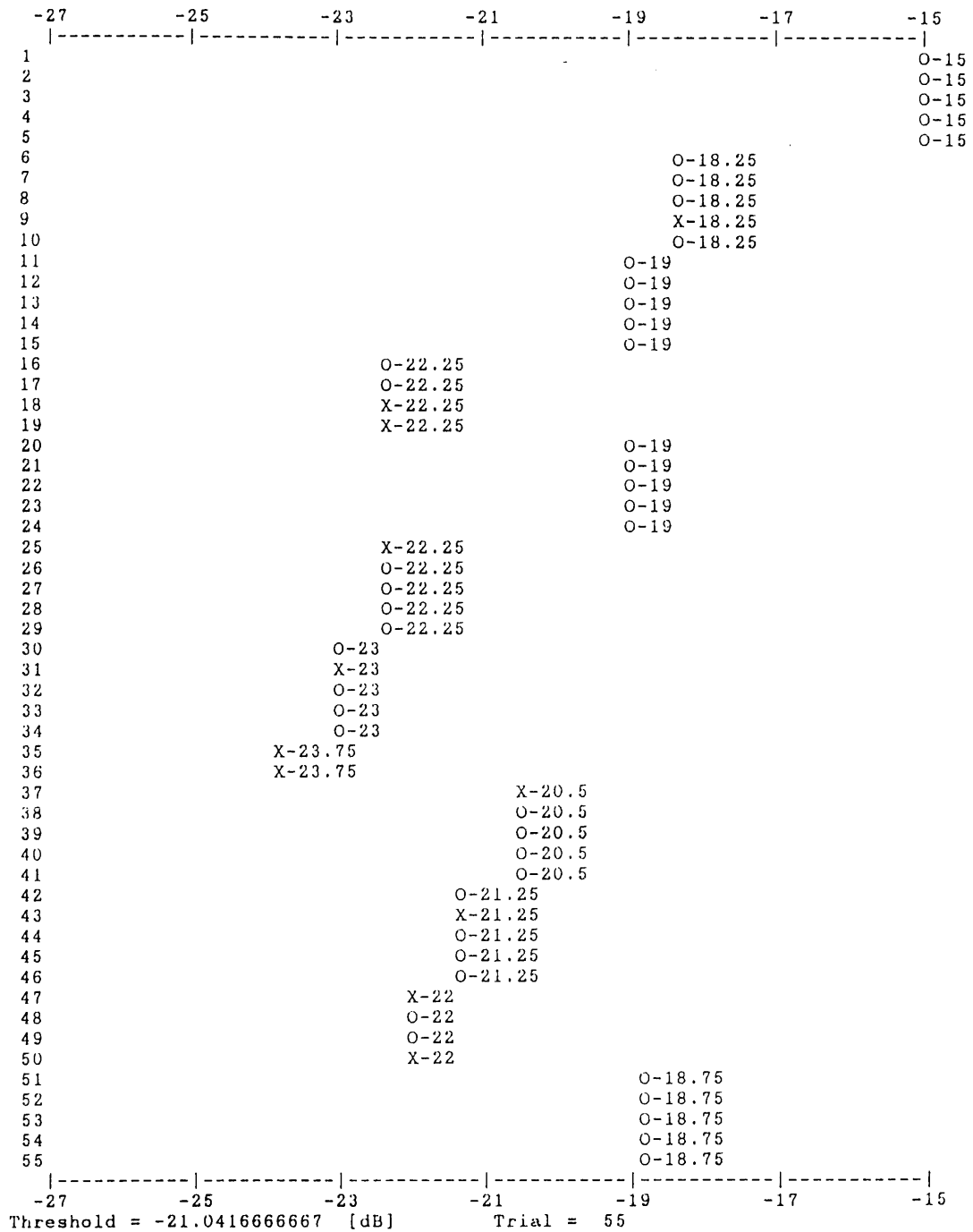


図 3.6: RASS 法の測定例 (横軸は刺激レベル、縦軸は試行数、正応答を○、誤応答を×で示す)

測定を始める前に決定することは、

- 心理測定関数の傾きから、各正応答率パターンでの刺激レベルの変化幅を決定する。

- 最小ステップサイズを決定する (コンピュータによる実時間処理で刺激音を作成できない場合、あらかじめ刺激音のデータを作成する必要がある。その場合、最小ステップ幅を決定し、刺激音をあらかじめ作成する)。
- 測定を開始する閾値上の刺激レベルを決定する。
- 測定終了を決める刺激レベルの変化方向が逆転する個数を決定する。

である。

3.3 まとめ

RASS 法の特徴について、以下に記述する。

- 刺激レベルの変化方法のアルゴリズムが非常に単純である。実験前にステップ幅を決定するため、測定に要する処理が速い。
- 心理測定関数をロジスティック関数と仮定し、その傾き推定値を使用することで、測定条件のパラメータが簡単に決定できる。
- 被験者が応答の押し間違いをしても、同じ刺激レベルで 1 回までなら、次に呈示される刺激レベルが下がるため、ある程度 lapse に対応できる。
- 刺激レベルの変化が起こりやすく、また、チャンスレベル付近の刺激レベルでは、比較的分かりやすい刺激レベルが呈示される (50% 以下の確率事象が出る確率が高く、その場合、ステップ幅も大きい)。

第 4 章

モンテ・カルロ・シミュレーションによる適応法の効率比較

4.1 はじめに

第 3 章で、新しく開発した RASS 法の紹介をした。RASS 法の主な特徴として、測定で設定するパラメータが簡単に決定され、刺激レベルの変化アルゴリズムが単純であるなどが挙げられる。第 2 章で、恒常法は聴取訓練を受けていない被験者群では、閾値の安定性が悪く、内観報告からも適応法に較べ良い結果が得られなかった。そこで、従来からある適応法 (PEST 法と Best PEST 法) と RASS 法を用いて、本章では、まず、コンピュータ・シミュレーションで効率の比較を行ない、様々な測定条件下での各測定法の効率を調べる。そして、次章で聴取実験により、実験効率と被験者の負担についても検討を行なう。

シミュレーションで効率を比較する利点は、真値 (閾値、心理測定関数の傾き) があらかじめ分かっていることである。シミュレーションで多数回の測定を行なえば、その求めた値の平均値と閾値 (真値) との差が分かる。その差をバイアスと呼ぶ。また、求めた値にばらつきが生じる。そのばらつきを標準偏差で表す。この二つのパラメータと試行数で効率を評価する。同じ試行数なら、バイアスと標準偏差が共に小さいほど効率が良い。

測定法の効率をシミュレーションで比較する場合、一般的に、モンテ・カルロ・シミュレーションが使用される。まず、心理測定関数を仮定する。心理測定関数が決まれば、刺激レベルと正応答率との対応関係が分かる。つまり、ある刺激レベルで刺激を呈示した場合、その刺激レベルでの正応答率が分かる。例えば、2IFC 法を用い正応答率が 80% の刺激レベルで、試行を行なった場合、80% の確率で正応答が得られ、20% の確率で誤応答が得られる。このことを利用して、シミュレーションを

行なう。応答結果は0~1の一様乱数で模擬し、1試行毎に乱数を発生させ、その刺激レベルでの心理測定関数の正応答率以下の値が出た場合は正応答とし、その値以上が出た場合は誤応答とする。図4.1に2IFC(2者強制選択法)で心理測定関数にロジステック関数を使用した場合の例を示す。

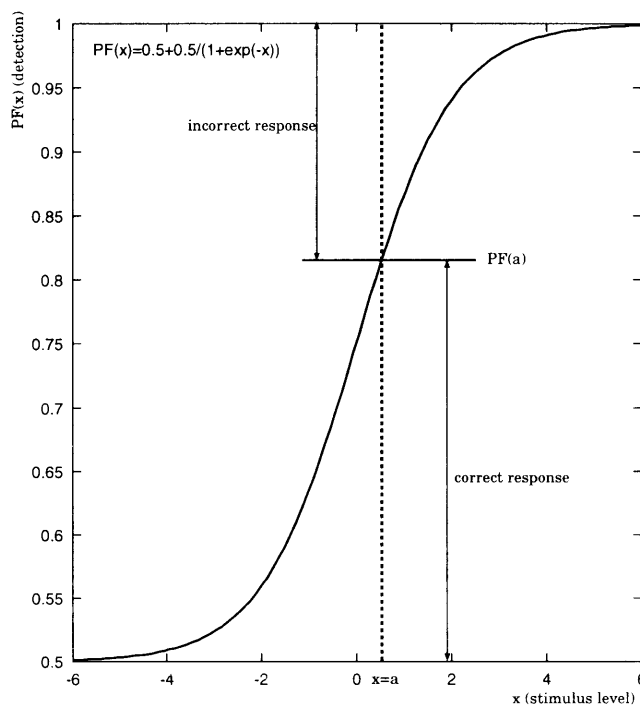


図 4.1: 心理測定関数にロジステック関数を使用した場合のモンテ・カルロ・シミュレーションでの正応答と誤応答の関係

本論文では、モンテ・カルロ・シミュレーションを用いて、適応法の効率を比較する。心理測定関数はロジステック関数(式4.1)を使用する。この関数はBest PEST法で使用されており、また、シミュレーションでも良く利用されている[7][12][21]。この関数のパラメータは、傾き値 a と閾値 b で決定され、真値 (true value) を傾き $a=1.0$ 、閾値 $b=0$ (logits) とする。

$$PF(x, a, b) = \frac{1}{2} + \frac{1}{2} \times \frac{1}{1 + \exp\{a \times (b - x)\}} \quad (4.1)$$

4.2 効率比較

4.2.1 共通するシミュレーション条件

PEST 法、Best PEST 法、RASS 法の 3 種類の適応法を用い、効率の比較をする。そこで、効率 (バイアス、標準偏差、試行数) を以下の点について検討する。

1. 刺激レンジが効率に及ぼす影響。
2. 心理測定関数の傾きが効率に及ぼす影響。
3. Best PEST 法において、閾値が刺激レンジの中心からずれた場合、効率に及ぼす影響。
4. 刺激レンジ及び刺激レンジ内での閾値が存在する位置との関係をランダムに変更した場合に、効率に及ぼす影響。

である。

心理測定関数は、ロジスティック関数 (式 4.1) を使用し、真値を傾き $a=1.0$ 、閾値 $b=0$ (logits) とした。そして、2IFC 法で 75% の正応答率の閾値を求めることにした。これは、Best PEST 法が 2IFC 法の場合、75% の正応答閾値しか求まらないため、これに合わせた。まず、各測定法で使用する最小ステップサイズを 0.125 (logits) に決定した。この章では、刺激レンジ (刺激レベルの上限と下限の幅) を 4~16 (logits) まで 1 (logits) 刻みで細かく変更し、測定を開始する刺激レベルが効率に及ぼす影響などを調べる。その理由として、測定を開始する刺激レベルが変更されることにより、測定法の効率がどのように変化するかを調べるためである。今回のシミュレーションでは、真値から最小ステップサイズ分ずれた値が求まった場合、そのずれが 75% 正応答率から 2% 以内とした。通常、多数回のシミュレーションを行なうと、バイアスは 0 ではないため、真値からずれることが予想できる。そのずれは、最小ステップサイズが大きな値だと、その結果、真値 (75% 閾値) からのずれも大きくなる。また、Best PEST 法はあらかじめ刺激レンジを等分割し、各々の刺激レベルを測定前に決定する。よって、刺激レンジを最小ステップサイズで割った数が整数となる必要がある。正の方向に真値から 2% の最小ステップサイズ分ずれて求まった場合、その値は、77% (2% の閾値からのずれ) で 0.16 (logits) となる。しかし、この値を刺激レンジで割ると整数にならない。よって、2% 以内で、最も近い値である 0.125 (logits) を使用した。なお、最小ステップサイズが 0.125 (logits) の場合、閾値 ($b=0$) から最小ステップサイズ分だけずれた値 ($x=0.125$) が求まった場合、その正応答率は 76.6%、つまり 1.6% ずれる。

また、PEST 法は stopping ステップサイズを一定にし、最大ステップサイズのみ変更した場合、測定条件が少ない。そこで、PEST 法のみ、最小ステップ幅を 0.125 と 0.25(logits) の 2 種類を用いた。

RASS 法は 75.2% の閾値が求まるが、75% との差が $x=0.016$ (logits) で最小ステップサイズより小さいため、75.2% から 75% の補正は行わなかった。

各測定法で共通するシミュレーション条件は、

- PEST … 「最大ステップサイズ及び初期ステップサイズが 1.0(logits) で、stopping ステップサイズ = 0.125 及び 0.25(logits) の場合」と 「最大ステップサイズ及び初期ステップサイズが 2.0(logits) で、stopping ステップサイズ = 0.125 及び 0.25(logits) の場合」。W(deviation limit) は、全ての場合において、1.0(推奨値) を使用。
- Best PEST 法…ステップサイズ = 0.125(logits)。試行数を 20~100 試行まで 10 試行区切り。
- RASS 法…最小ステップサイズ = 0.125(logits)。刺激レベルの変化方向の逆転する山、谷の個数が 2~10 で 1 個区切り。

である。

刺激レンジを設定しない PEST 法と RASS 法は、測定を開始する刺激レベルを刺激レンジの上端にした。シミュレーションのブロック (測定区切りの最小単位) 数は、5000 回 (1000 回毎に乱数の種を変更) とした。なお、乱数には Numerical Recipes in C[22] の ran2 を使用した。この乱数は L'Ecuyer の長周期の乱数生成法と Bays-Durham の切り混ぜに安全機構をつけたものであり、コンピュータ演算の浮動小数点の精度の限界を考慮すれば、ほぼ完全な乱数である。

4.2.2 刺激レンジが効率に及ぼす影響

4.2.2.1 シミュレーション条件

測定を行なう上で刺激レンジを決定することは重要である。特定の刺激レベルから測定を開始した場合にのみ、効率が良い測定法があれば、その測定法は測定を開始する刺激レベルとおよその閾値が分かっている場合にしか使用できない。そこで、刺激レンジを 4~16(logits) まで 1(logits) 刻みで調べることにした (PEST 法と RASS 法の測定開始の刺激レベルは、2~8(logits) で 0.5(logits) 刻みになる)。その他のシミュレーション条件は、共通する実験条件 (4.2.1) と同じである。

PEST 法と RASS 法は、徐々に閾値に近付いていく刺激レベルの変化アルゴリズムの特性のため、刺激レンジが広くなるにつれ、試行数が増加することが推測できる。

4.2.2.2 結果

横軸に試行数、縦軸に閾値のバイアス及び標準偏差を取った各適応法の結果の一部を図 4.2~4.9 に示す (その他の結果は付録 A に記す)。ただし、PEST 法及び RASS 法は平均試行数でバイアス及び標準偏差を示した。また、PEST 法及び RASS 法の試行数の平均値及び標準偏差の結果の一部を表 4.1~4.8 に示す (その他の結果は付録 A に記す)。

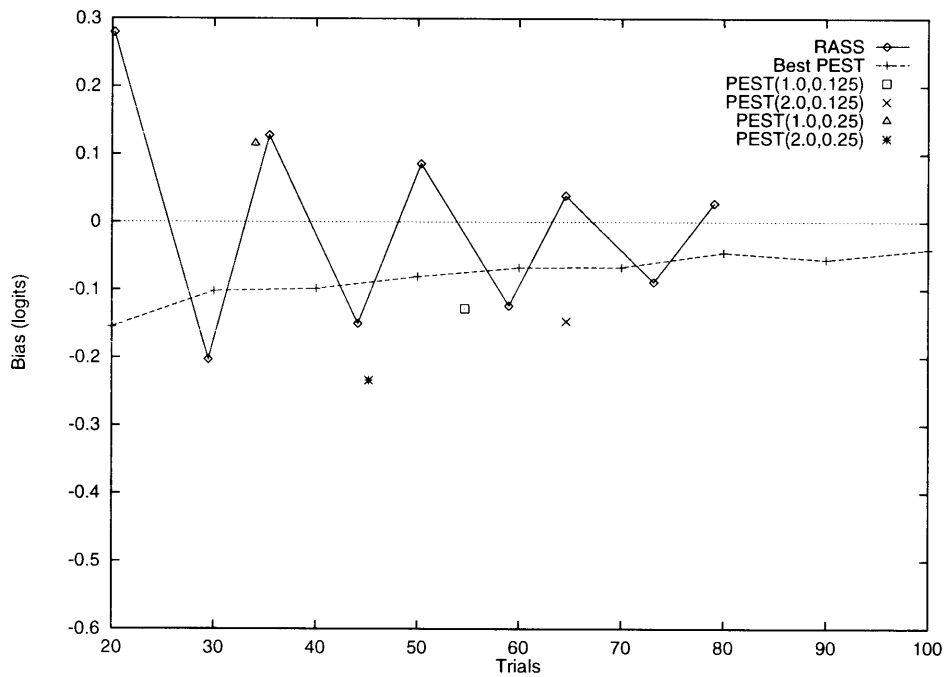


図 4.2: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 4(logits) である。)

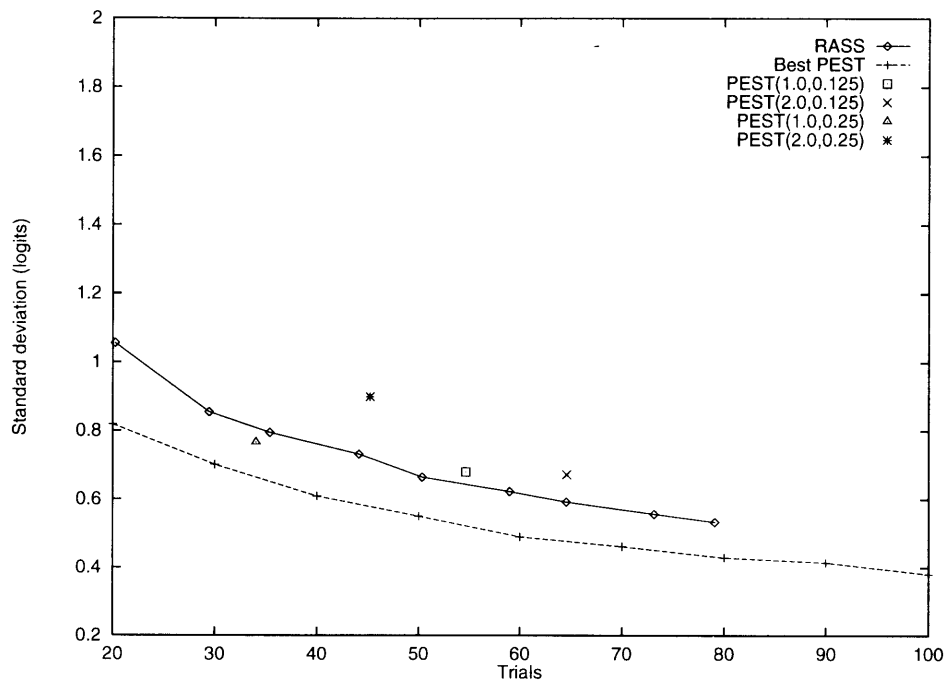


図 4.3: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 4(logits) である。)

表 4.1: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 4(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 20.2 | 29.4 | 35.4 | 44.1 | 50.3 | 59.0 | 65.5 | 73.2 | 79.1 |
| Standard deviation | 6.8 | 8.5 | 9.1 | 10.5 | 11.0 | 12.3 | 12.5 | 13.7 | 14.1 |

表 4.2: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 4(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 54.7 | 27.4 |
| PEST(2.0,0.125) | 64.6 | 33.8 |
| PEST(1.0,0.25) | 34.0 | 19.5 |
| PEST(2.0,0.25) | 45.2 | 22.9 |

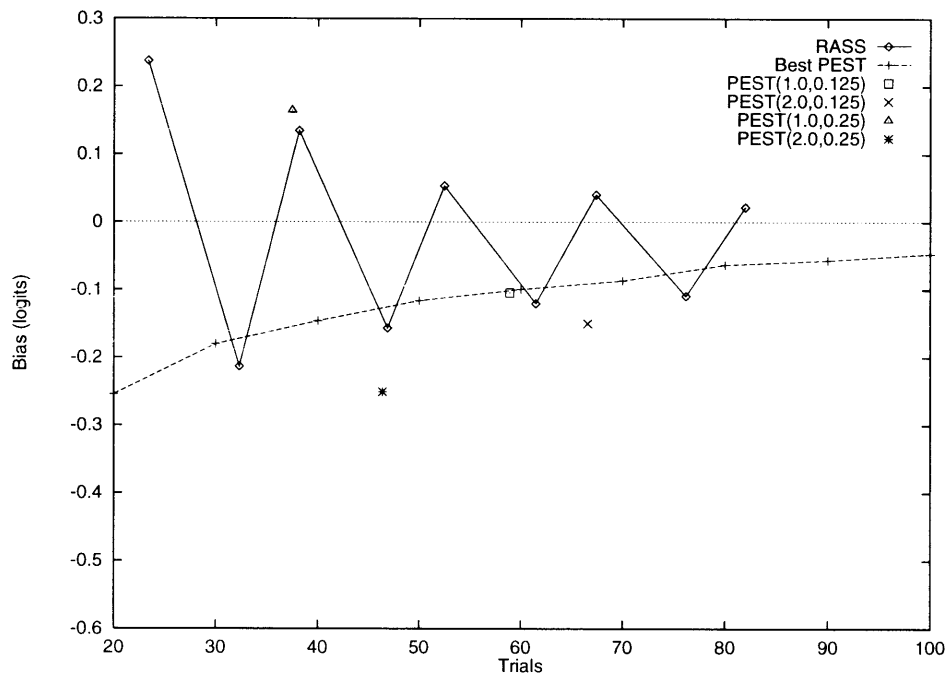


図 4.4: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 6(logits) である。)

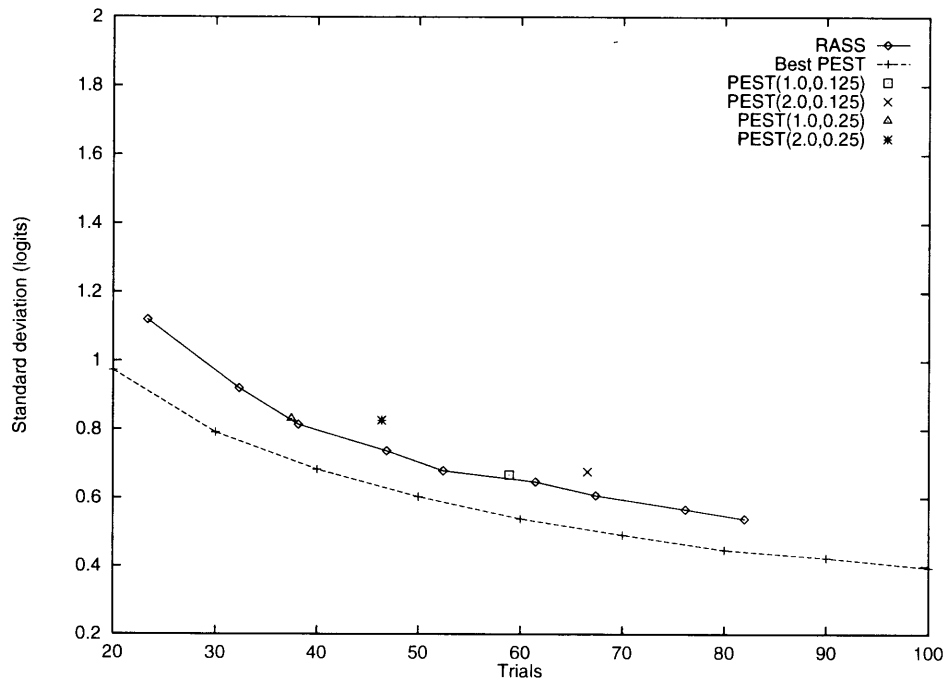


図 4.5: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 6(logits) である。)

表 4.3: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 6(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 23.3 | 32.3 | 38.1 | 46.9 | 52.4 | 61.5 | 67.4 | 76.2 | 82.0 |
| Standard deviation | 7.1 | 8.4 | 9.2 | 10.7 | 10.9 | 12.2 | 12.4 | 13.9 | 14.0 |

表 4.4: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 6(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 58.9 | 27.0 |
| PEST(2.0,0.125) | 66.6 | 34.2 |
| PEST(1.0,0.25) | 37.5 | 19.0 |
| PEST(2.0,0.25) | 46.4 | 22.9 |

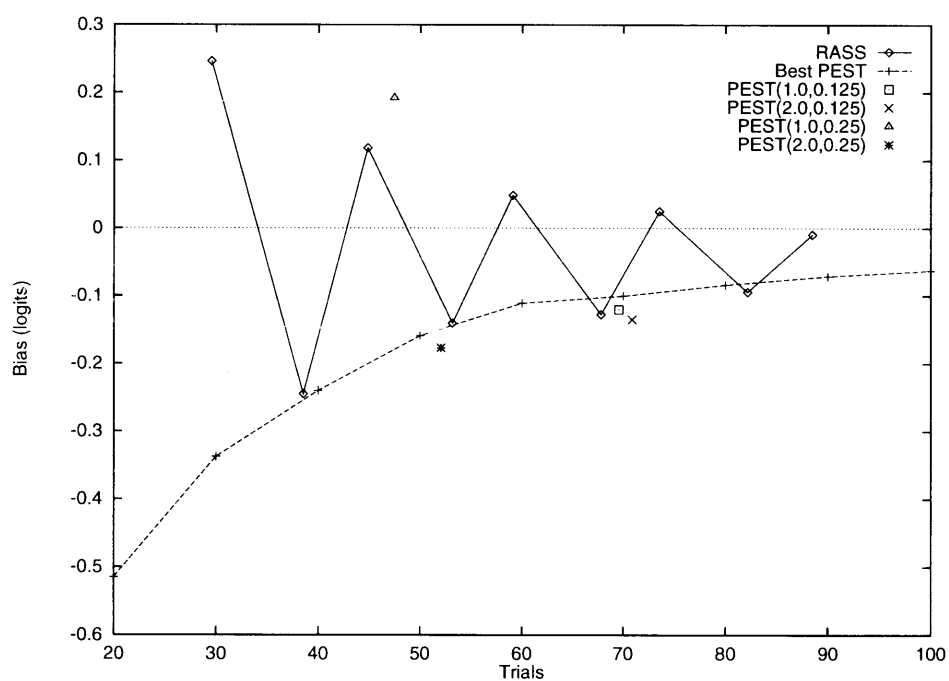


図 4.6: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 11(logits) である。)

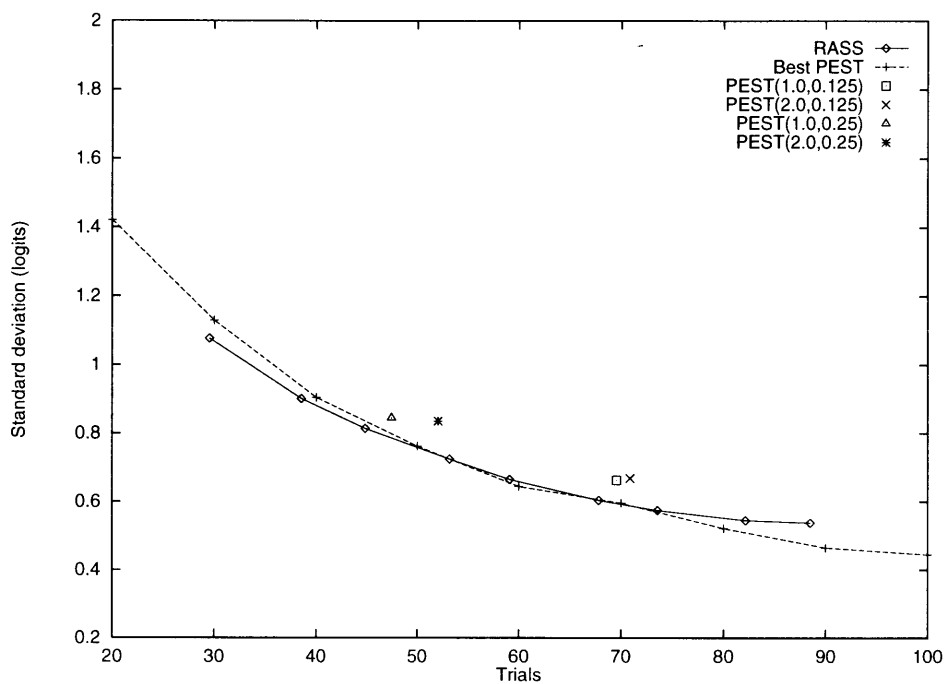


図 4.7: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 11(logits) である。)

表 4.5: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 11(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 29.5 | 38.5 | 44.9 | 53.2 | 59.0 | 67.8 | 73.6 | 82.1 | 88.5 |
| Standard deviation | 7.2 | 8.6 | 9.4 | 10.7 | 11.1 | 12.5 | 12.7 | 13.6 | 14.2 |

表 4.6: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 11(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 69.6 | 27.7 |
| PEST(2.0,0.125) | 70.9 | 34.6 |
| PEST(1.0,0.25) | 47.4 | 19.1 |
| PEST(2.0,0.25) | 52.1 | 24.1 |

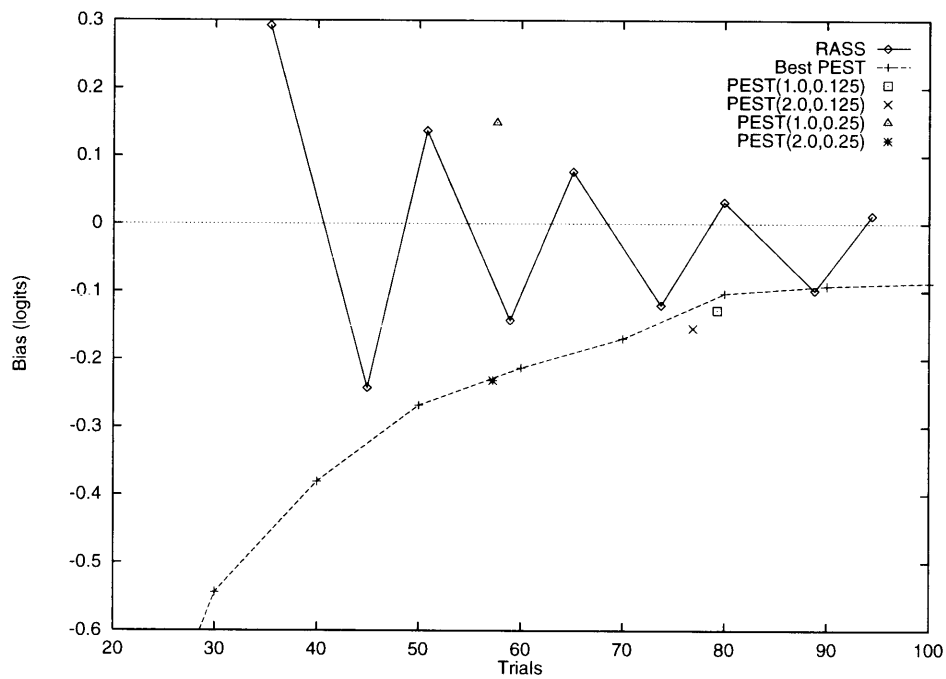


図 4.8: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 16(logits) である。)

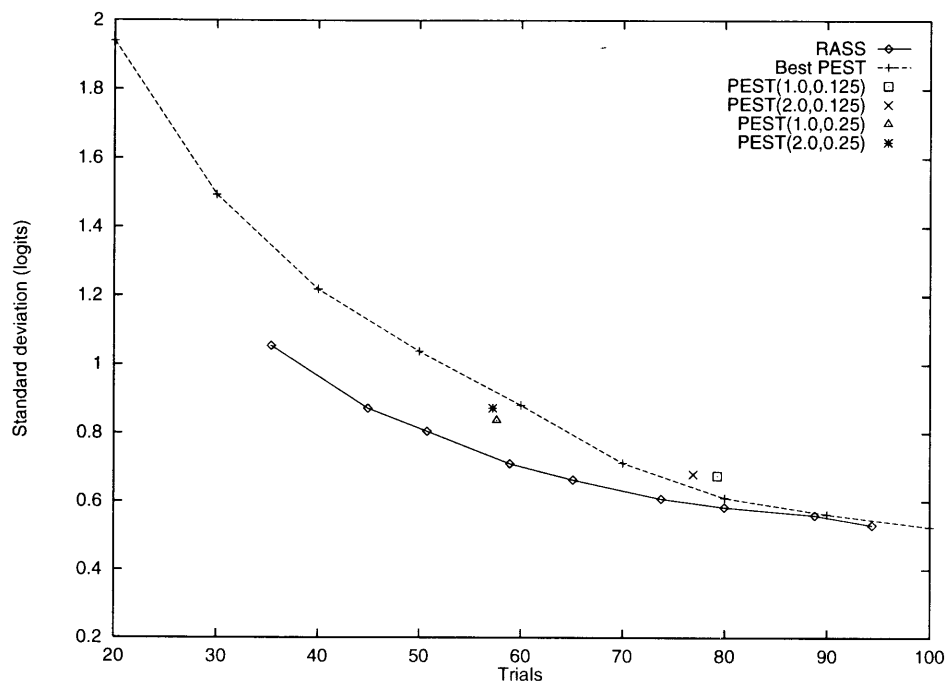


図 4.9: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 16(logits) である。)

表 4.7: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 16(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 35.4 | 44.9 | 50.8 | 58.9 | 65.1 | 73.8 | 80.0 | 89.0 | 94.4 |
| Standard deviation | 6.9 | 8.6 | 9.3 | 10.6 | 11.0 | 12.2 | 12.8 | 13.8 | 14.2 |

表 4.8: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 16(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 79.3 | 27.2 |
| PEST(2.0,0.125) | 76.9 | 34.5 |
| PEST(1.0,0.25) | 57.6 | 19.2 |
| PEST(2.0,0.25) | 57.2 | 23.6 |

刺激レンジが狭い場合 (図 4.3)、標準偏差は Best PEST 法が最も小さな値をとる。これは、他の適応法は、刺激レンジ外の刺激レベルの呈示もあるが、Best PEST 法は設定した刺激レンジ外の刺激の呈示がないためであると思われる。バイアスは、RASS 法が刺激レンジが 6(logits) 以上の場合 (図 4.4)、最も小さい値を取る (ただし、刺激レベルが逆転する折り返し数が偶数の場合である)。刺激レンジが増加するにつれ、RASS 法の標準偏差の値は変化しないが、Best PEST 法は徐々に大きな値になる。そして、刺激レンジが 11(logits) より大きくなるにつれ、RASS 法が Best PEST 法と比較して標準偏差が小さな値を取る (図 4.7)。さらに、Best PEST 法は刺激レンジが増加するとバイアスも大きな値になるが、RASS 法はあまり変化しない (図 4.2、図 4.8)。このことは、実験を行なう上で、閾値上の刺激レベルなら、どの刺激レベルから測定を開始しても良いという利点があり、予備実験で被験者の閾値を測定する場合、明らかに分かる刺激レベルから測定を開始しても、試行数が若干多くなるだけで、バイアスと標準偏差の値は変わらない。これに対し、Best PEST 法は刺激レンジが広くなるにつれ、バイアスと標準偏差は大きな値を取り、効率が悪くなるため、刺激レンジを広く設定することができない。また、Best PEST 法は、被験者の閾値が設定した刺激レンジ内に入っていない場合、その試行が無駄になるという欠点もある。

PEST 法と RASS 法は刺激レンジが増加するにつれ、平均試行数が徐々に増加する (表 4.1~4.8)。これは、刺激レベルのアルゴリズム上の問題点であるが、刺激レンジが 4 と 16(logits) の差が約 16 試行であり、あまり問題にならないと思われる。しかし、試行数の標準偏差は、刺激レンジが増加しても変化しない。PEST 法と RASS 法の試行数の標準偏差を比較すると、RASS 法が PEST 法の標準偏差の約半分の値ぐらいである (表 4.7 と表 4.8 など)。

PEST 法は stopping ステップサイズが 0.125(logits) の場合、試行数も増え、バイアス及び標準偏差が最も悪いことが分かった (表 4.8 など)。PEST 法の最大ステップサイズが 1.0(logits)、stopping ステップサイズが 0.25(logits) の場合のみ、バイアスが正の値を取った。この理由は、ステップ幅が 1.0、0.5、0.25(logits) の 3 段階で終了したためであると思われる (その他の PEST 法の条件では、4 段階以上である)。

Best PEST 法は刺激レンジが増加するにつれ、バイアスと標準偏差が共に大きくなる。

RASS 法は山・谷の折り返しが偶数の場合にバイアスが小さい。そのため、測定は偶数個で終了すべきである。また、バイアス及び標準偏差は刺激レンジが増加しても、試行数が増加するだけで、あまり変化しない。

4.2.3 心理測定関数の傾き推定値が効率に及ぼす影響

4.2.3.1 シミュレーション条件

PEST 法は心理測定関数を仮定しないが、Best PEST 法及び RASS 法では、心理測定関数の傾き推定値を使用する。そこで、心理測定関数の傾き推定値が真値 ($a=1.0$) と mismatching の場合について、シミュレーションを行い、効率を調べる。傾き推定値は、 $a=0.5$ 、0.75、1.0、1.25、1.5、1.75、2.0 を使用し、傾き推定値が overestimate と underestimate の場合の効率を調べる。第 2 章の恒常法の実験結果より、被験者毎の心理測定関数の傾き値が 0.6 で、その標準偏差は約 0.2 であった。よって、今回のシミュレーション条件の傾き推定値が真値の 1/2 倍と 2 倍は、非常に極端な場合のシミュレーションである。

刺激レンジは、ロジスティック関数の 51%~99% に近い 8(logits) を使用した。その他の測定法毎のシミュレーション条件は、共通するシミュレーション条件 (4.2.1) を使用した。

4.2.3.2 結果

横軸に試行数、縦軸にバイアス及び標準偏差を取り、結果を図 4.10~4.15 に示す。

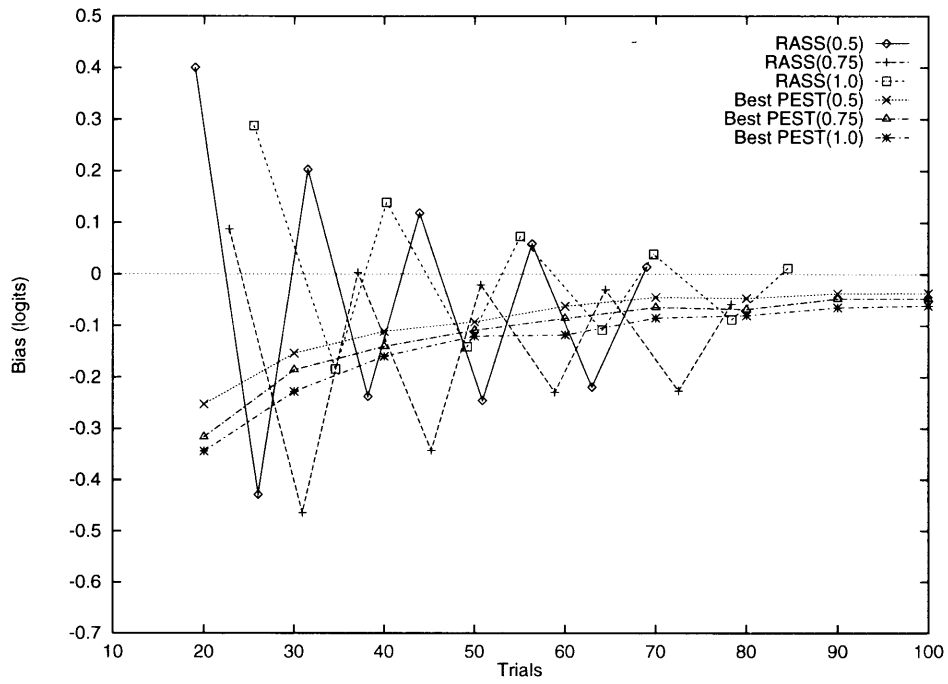


図 4.10: 心理測定関数の傾き推定値を小さく見積もられた場合 (真値の 0.5~1.0 倍) の影響 (RASS 法と Best PEST 法で閾値推定のバイアスの結果を示す。RASS(x),BP(x) は真値 ($a=1.0$) の x 倍の値を使用した結果である。)

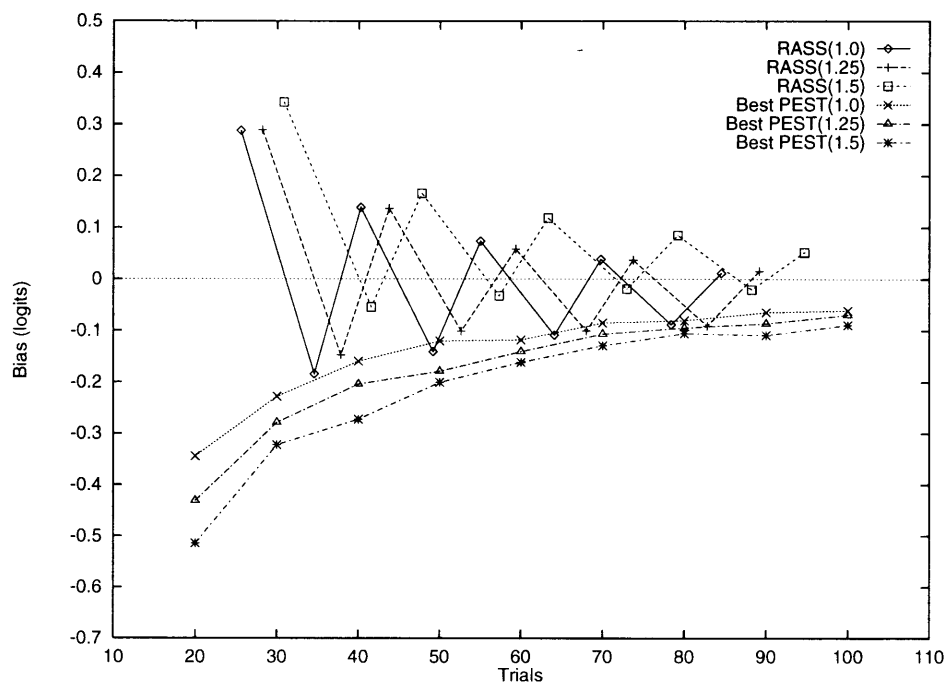


図 4.11: 心理測定関数の傾き推定値を大きく見積もった場合 (真値の 1.0~1.5 倍) の影響 (RASS 法と Best PEST 法で閾値推定のバイアスの結果を示す。RASS(x),BP(x) は真値 ($a=1.0$) の x 倍の値を使用した結果である。)

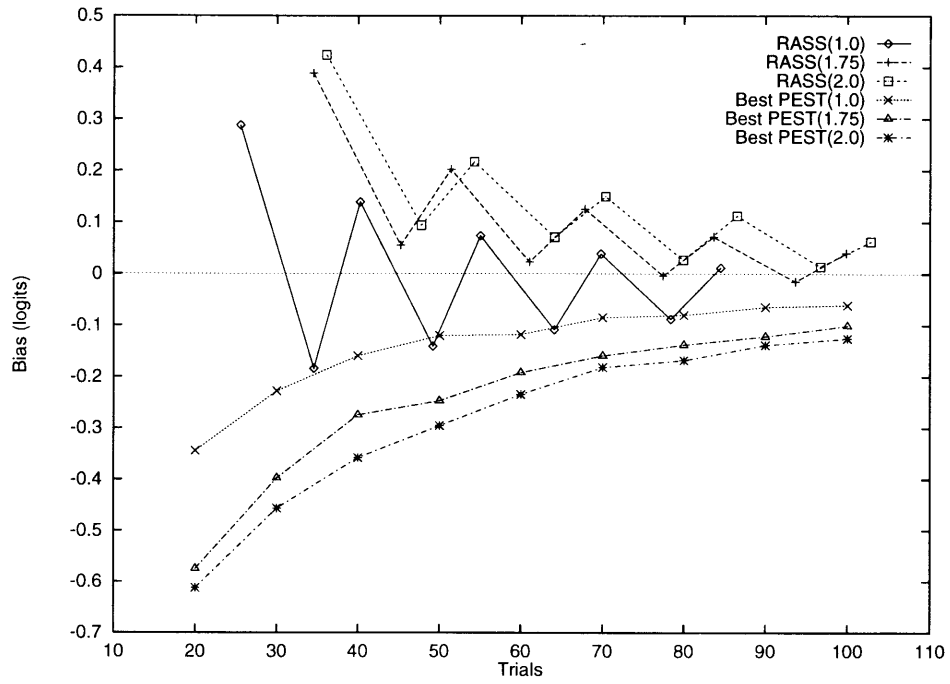


図 4.12: 心理測定関数の傾き推定値を大きく見積もった場合 (真値の 1.75~2.0 倍) の影響 (RASS 法と Best PEST 法で閾値推定のバイアスの結果を示す。RASS(x),BP(x) は真値 (a=1.0) の x 倍の値を使用した結果である。)

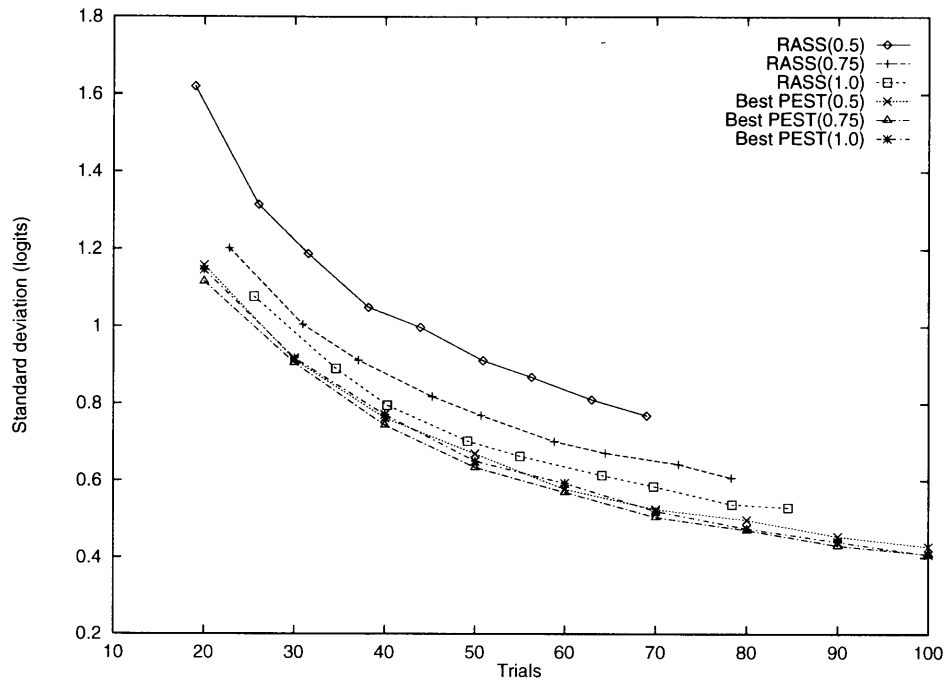


図 4.13: 心理測定関数の傾き推定値を小さく見積もった場合 (真値の 0.5~1.0 倍) の影響 (RASS 法と Best PEST 法で閾値推定の標準偏差の結果を示す。RASS(x),BP(x) は真値 (a=1.0) の x 倍の値を使用した結果である。)

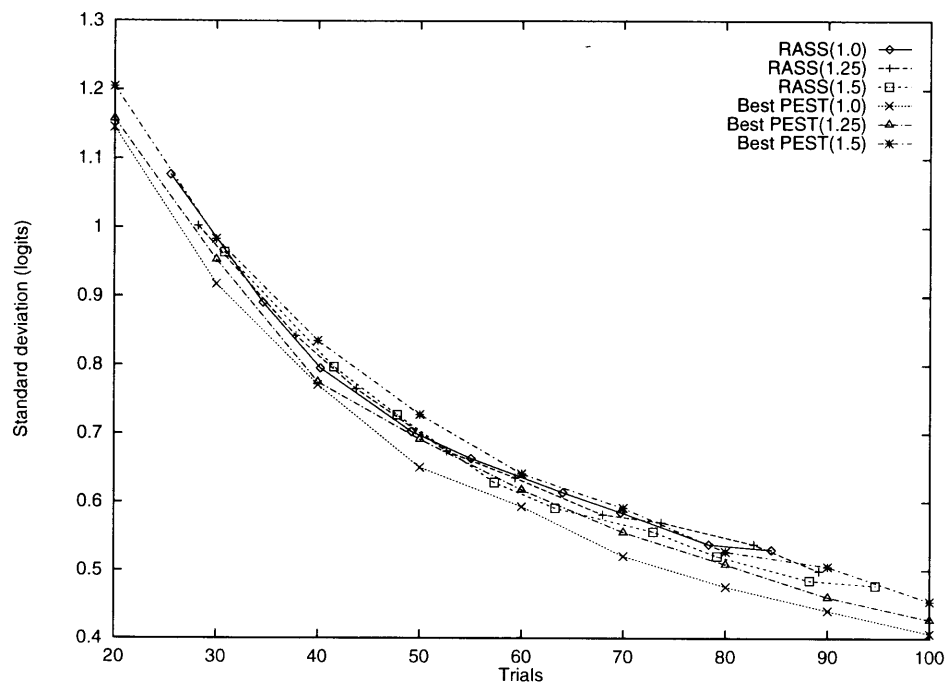


図 4.14: 心理測定関数の傾き推定値を大きく見積もった場合 (真値の 1.0~1.5 倍) の影響 (RASS 法と Best PEST 法で閾値推定の標準偏差の結果を示す。RASS(x),BP(x) は真値 ($a=1.0$) の x 倍の値を使用した結果である。)

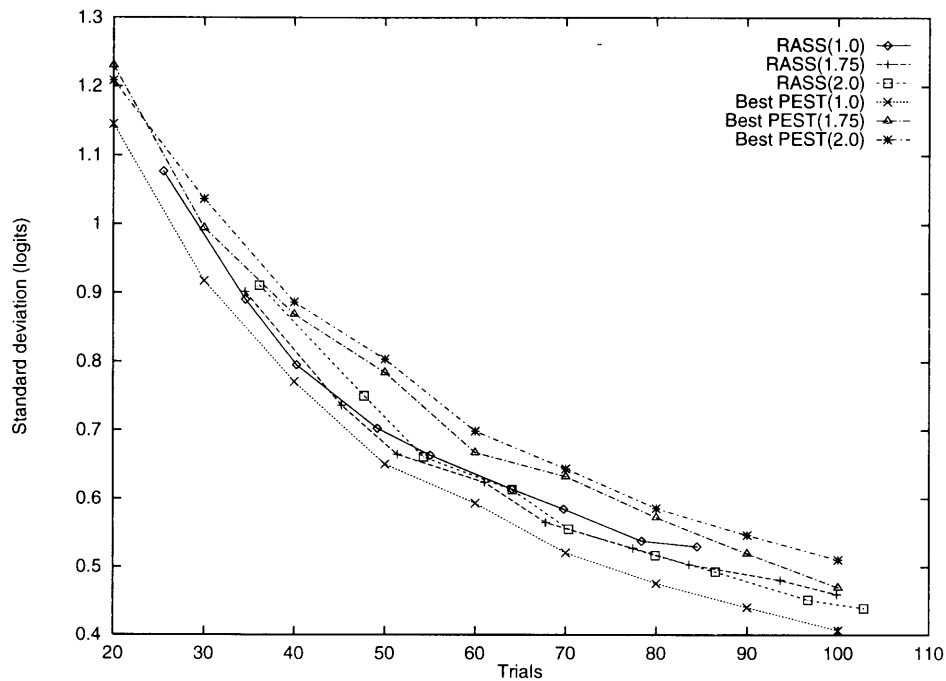


図 4.15: 心理測定関数の傾き推定値が大きく見積もった場合 (真値の 1.75~2.0 倍) の影響 (RASS 法と Best PEST 法で閾値推定の標準偏差の結果を示す。RASS(x),BP(x) は真値 ($a=1.0$) の x 倍の値を使用した結果である。)

Best PEST 法は心理測定関数の傾き推定値が真値より小さい場合、バイアスは小さな値になり、標準偏差はほぼ同じであるため、効率が良くなる。傾き推定値が真値より大きな値になるにつれ、バイアス及び標準偏差も大きくなり、効率が悪くなる。

RASS 法は傾き推定値が真値より小さい場合、平均試行数が減少し、バイアスは平均試行数が減少したことを考慮にいと、ほぼ同じであるが、標準偏差は大きな値になり、効率が悪くなる。傾き推定値が真値より少し大きい場合、バイアスが最も小さな値になる。それ以上の傾き推定値ではバイアスは大きな値になるが、標準偏差はほぼ同じである。

4.2.4 Best PEST 法において、刺激レンジ内の閾値が刺激レンジの中心からずれた場合、効率に及ぼす影響について

4.2.4.1 シミュレーション条件

PEST 法と RASS 法は刺激レンジを設定しないため、刺激レンジ内の閾値の場所の影響を受けない。なぜなら、設定した閾値と測定を開始する刺激レベルまでの距

離が関係するためである。前節まで、Best PEST 法は刺激レンジの中心にのみ閾値を設定し、シミュレーションを行なった。そこで、Best PEST 法で刺激レンジ内の閾値の場所をずらした場合の効率を調べる。シミュレーション条件は、刺激レンジを 8(logits) に固定し、閾値を 0.5(logits) ステップで -2~2(logits) まで変更した。その他の条件は、共通するシミュレーション条件 (4.2.1) と同じである。

4.2.4.2 結果

横軸に試行数、縦軸に閾値のバイアス及び標準偏差を取り、図 4.16、4.17に結果を示す。

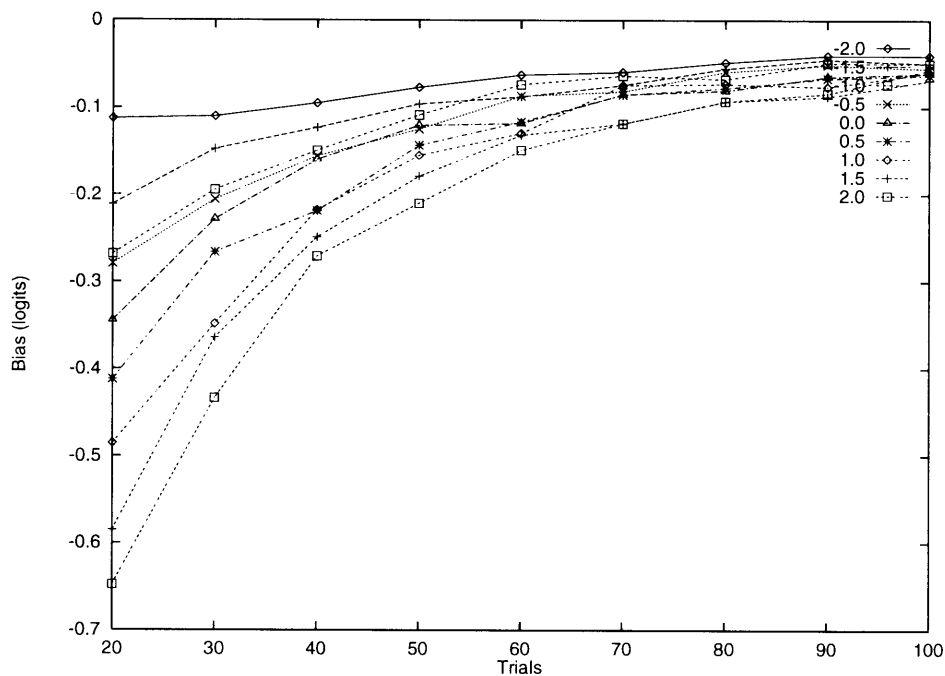


図 4.16: 刺激レンジ内の閾値が刺激レンジの中心からずれた場合、Best PEST 法の効率に及ぼす影響についてのバイアスの結果

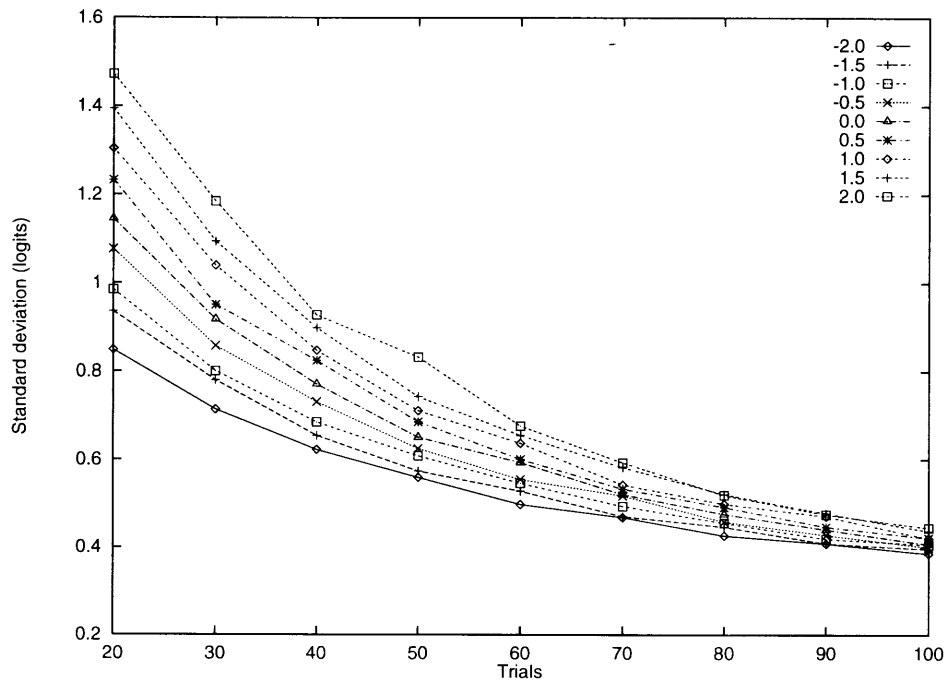


図 4.17: 刺激レンジ内の閾値が刺激レンジの中心からずれた場合、Best PEST 法の効率に及ぼす影響についての標準偏差の結果

Best PEST 法は、閾値が刺激レンジの中心より、低い刺激レベルにある場合、バイアス及び標準偏差の値が小さく、効率が良い。また、閾値が刺激レンジの中心より、高い刺激レベルにある場合、バイアス及び標準偏差が悪くなる。このことは、Best PEST 法の最尤推定のアルゴリズムが関係していると思われる。

4.2.5 刺激レンジ及び刺激レンジ内での閾値が存在する位置との関係をランダムに変更した場合に、効率に及ぼす影響

4.2.5.1 シミュレーション条件

3種類の適応法を用いたシミュレーションの効率比較では、絶えず閾値が刺激レンジの中心に位置していた。そこで、刺激レンジ及び閾値の位置をランダムに選び、シミュレーションを行う。実験では閾値の位置はある程度の刺激レベルの範囲でしか、分からない。そこで、1ブロック毎に刺激レンジを 6.0~16.0(logits) 内で 0.25(logits) ステップでランダムに決定し、閾値も -2.0~2.0(logits) 内で 0.125(logits) ステップでランダムに決定した。それぞれの値を決定後、1ブロックの測定を行なう(測定中の閾値は固定している)。そして、1ブロックの測定が終了すると、再度、閾値と刺激レンジをランダムに変更した。そのランダムにした閾値と刺激レンジの分布は

一様分布である。今回のシミュレーション条件では、ブロック数を 10000 回 (1000 回毎に乱数の種を変更) と多くとることで、閾値及び刺激レンジをランダムに選ぶことによる測定法毎の条件の偏りをなくす。その他のシミュレーション条件は、共通するシミュレーション条件 (4.2.1) と同じである。

4.2.5.2 結果

横軸に試行数、縦軸に閾値のバイアス及び標準偏差を取り、図 4.18、4.19 に結果を示す。また、RASS 法と PEST 法の平均の試行数および標準偏差を表 4.9、4.10 に示す。

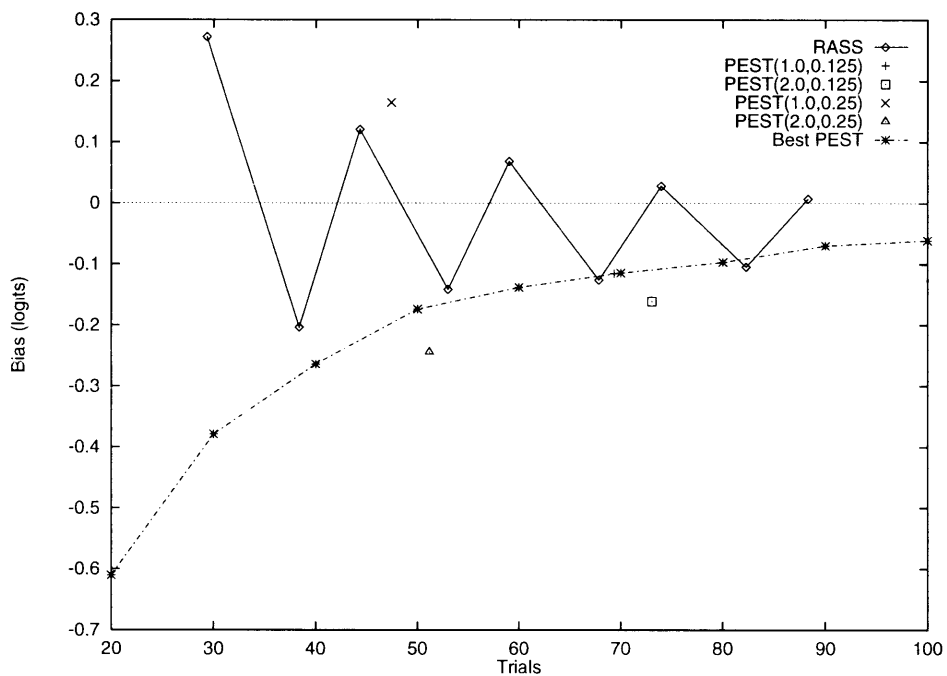


図 4.18: 閾値と刺激レンジをそれぞれランダムに変更した場合のバイアスの結果 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。)

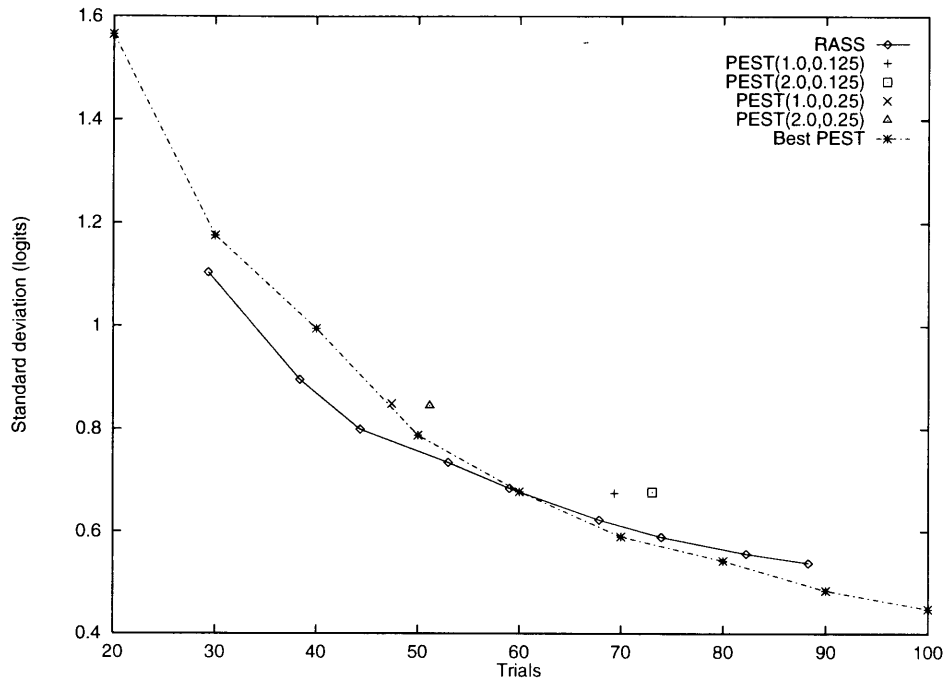


図 4.19: 閾値と刺激レンジをそれぞれランダムに変更した場合の標準偏差の結果 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。)

表 4.9: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 29.3 | 38.4 | 44.3 | 53.0 | 59.0 | 67.9 | 74.0 | 82.3 | 88.3 |
| Standard deviation | 8.5 | 10.0 | 10.0 | 11.4 | 11.8 | 13.1 | 13.5 | 14.4 | 14.8 |

表 4.10: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズである。)

| 条件 | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 69.3 | 28.5 |
| PEST(2.0,0.125) | 73.1 | 35.5 |
| PEST(1.0,0.25) | 47.4 | 20.6 |
| PEST(2.0,0.25) | 51.2 | 23.3 |

バイアスは RASS 法が最も小さな値をとる。PEST 法の stopping size が 0.25(logits) で、最大ステップサイズが 1.0(logits) の場合、Best PEST 法とほぼ同じ効率である。その他の PEST 法の条件では、他の適応法と比較すると、効率が悪いことが分かった。標準偏差は、RASS 法が 60 試行までは最も小さく、それ以上の試行数では、Best PEST 法が最も小さな値をとる。しかし、60 試行以上で RASS 法と Best PEST 法の標準偏差を比較するとほぼ同じ値を取る。そのため、バイアスの結果も考慮すると、RASS 法が最も効率が良いことが分かった。

試行数の標準偏差を RASS 法と PEST 法で比較すると、RASS 法は PEST 法の約 1/2 以下の値である。そのため、PEST 法と比較すると、RASS 法が測定終了までの時間が予測しやすく、実験計画を立てやすいことが分かった。

4.3 まとめ

モンテ・カルロ・シミュレーションで各測定法の効率を調べた。その結果、以下のことが分かった。

- PEST 法は、ほとんどの条件で、適応法の中で最も効率が悪かった。また、試行数の標準偏差が大きく、測定に要する時間が予測しにくいため、実験計画を立てにくいことが分かった。
- Best PEST 法は、刺激レンジが狭い場合に効率が良いことが分かった。そして、刺激レンジが広くなるにつれ、効率が悪くなる。これは、他の適応法と比較すると、刺激レンジ外の刺激が呈示されないためであると思われる。また、心理測定関数の傾き推定値が、真値より小さな値の時に効率が良いことが分かった。

た。また、閾値が刺激レンジの中心より、低い刺激レベルにずれている場合に効率が良いことが分かった。

- RASS 法は、刺激レベルの変化方向が逆転する山・谷の折り返し数が偶数の場合にバイアスが小さいことが分かった。刺激レンジが広くなるにつれ、他の適応法と比較すると効率が良くなる。測定が終了する試行数は、あらかじめ決まっていなくても、PEST 法と比較すると、試行数の標準偏差が小さく、測定に要する時間が予測しやすいため、実験計画を立てやすい。また、心理測定関数の傾き推定値が、真値より少し大きな値の時に効率が良い。そして、さまざまな条件の時に、他の適応法と比較すると、最も効率が良かった。

ここで、RASS 法の特徴について、この章で分かったことを記述する。

- 2IFC 法で効率が良い。
- 測定試行数の標準偏差が小さいため、測定に要する時間が予測しやすく、実験計画を立てやすい。
- 刺激レベルの変化方向が逆転する山・谷の折り返し数が偶数の場合にバイアスが小さい。
- 刺激レンジが広い時に、効率が良い。
- さまざまな条件の時に、効率が良い。

また、本論文で使用した適応法の特徴を表 4.11 にまとめてある。

表 4.11: 本論文で使用する各測定法の特徴

| | PEST | Best PEST | RASS |
|---------|---|--------------------------------|--|
| 心理測定関数 | 使用しない | ロジスティック関数の傾き値を使用する | ロジスティック関数の傾き値を使用する |
| 試行数 | あらかじめ決められたステップサイズになるまで、測定を続ける。一定ではない。 | 実験者が決定するため一定である。 | 山・谷の折り返し数で決定されるため、一定ではない。しかし、PEST法に較べると標準偏差が小さい |
| 刺激レンジ | 設定しない | 設定する(刺激レンジ外の閾値に対応できない) | 設定しない |
| 測定法の条件 | 最大ステップサイズ、初期ステップサイズ、W(deviation limit)、stopping ステップサイズ | 心理測定関数の傾き値、刺激レンジ、最小ステップサイズ、試行数 | 心理測定関数の傾き値、山・谷の繰り返し数、最小ステップサイズ(実時間で刺激音が作成できない場合) |
| 求まる正応答率 | 任意 | 2IFC法の場合 75% | 2IFC法で 75.2% |
| アルゴリズム | 少し複雑 | 複雑 | 単純 |
| 処理時間 | 速い | 遅い | 速い |
| 呈示刺激 | 徐々に閾値に近づく | 急激に閾値付近を測定する | 徐々に閾値に近づく |

第 5 章

適応法の効率比較実験

5.1 はじめに

第 2 章で、恒常法と適応法の振幅変調検知閾の実験を聴取訓練を受けた被験者と受けていない被験者群に分けて行ない、それぞれの効率を比較をした。その結果、聴取訓練を受けていない被験者群では、恒常法は被験者毎の閾値の安定性が悪く、効率が悪いことが分かった。そして、Best PEST 法の効率が良いことが分かったが、被験者の内観報告からは PEST 法が他の測定法より、疲労感が少ないことが分かった。そこで、従来の測定法より、効率が良く、かつ被験者の負担が少ない測定法 (RASS 法) を新たに開発した。第 3 章では、RASS 法の理論及び特徴を述べた。また、第 4 章では、RASS 法と他の適応法 (PEST 法、Best PEST 法) の効率をシミュレーションで調べた。その結果、RASS 法は、刺激レンジ (刺激レベルの上限と下限の幅) が広がっても、試行数が若干増加するだけで、バイアスと標準偏差はあまり変わらない値を取り、他の適応法と比較すると、最も効率が良いことが分かった。そして、さまざまな条件でも RASS 法が最も効率が良いことが分かった。第 4 章でのシミュレーションの利点は、真値 (閾値、傾き) が分かっているため、測定法の違いによるバイアス (真値からのずれの平均値) 及び標準偏差の値を同じ試行数で比較することで、測定法の効率が比較できた。しかし、実際の実験で閾値が正確に測定できるかは、実験を行なわないと分からない。そこで、第 2 章と同様の聴取実験を行ない各適応法の効率、及び被験者への負担を比較する。被験者は、聴取訓練を受けた被験者と受けていない被験者群で分けて行なう。そして、内観報告を両被験者群から取り、被験者の負担について調べ、RASS 法の被験者への負担が他の適応法と比較して、少ないかどうかを調べる。

5.2 振幅変調検知閾実験

5.2.1 共通する実験条件

第2章と同様の実験内容である振幅変調検知閾の測定を行ない、適応法の効率を比較した。

使用した測定法は第4章のシミュレーションと同様に PEST 法、Best PEST 法、RASS 法の3種類の適応法である。

刺激音及び PEST 法と Best PEST 法の測定条件は第2章と同様である。新たに RASS 法が加わったため、RASS 法の実験条件を他の適応法と同じになるように決定する。第2章で、最小ステップサイズを 0.25dB としたため、RASS 法も同様に 0.25dB とした。そして、試行数を 50 試行前後にするため、第4章のシミュレーションの結果より、刺激レベルの変化方向が逆転する山・谷の折り返し数を 6 回とした。Best PEST 法及び RASS 法は心理測定関数を仮定するため、ロジスティック関数の傾きを第2章の実験と同様に $a=0.6$ (式 3.7) とした。ここで、聴取訓練を受けた被験者と受けていない被験者群で共通する実験条件は、

- PEST 法…初期ステップサイズ = 2dB、最大ステップサイズ = 2dB、 $W(\text{deviation limit})=1$ 、stopping ステップサイズ = 0.25dB。
- Best PEST 法… 49 刺激レベル (刺激レンジ = 12dB、0.25 dB step)、心理測定関数 (ロジスティック関数) の傾き $a = 0.6$ 、1 ブロック = 50 試行。
- RASS 法…心理測定関数 (ロジスティック関数) の傾き $a=0.6$ 、最小ステップサイズ = 0.25dB、山・谷の折り返し数 = 6。

である。

Best PEST 法のみ刺激レンジをあらかじめ設定する必要がある。そこで、刺激レンジの中心が予備実験より予想される閾値になるように設定し、PEST 法及び RASS 法は予備実験の結果から求めた閾値の約 6dB 上から実験を行った。

5.2.2 聴取訓練を受けた被験者を用いた実験

5.2.2.1 実験条件

被験者は聴取訓練を受けた 6 人 (著者を含む) の九州芸術工科大学音響設計学科の学生 (21 から 27 歳) である [17][18][19]。

1 日に特定の測定法のみを用いると実験日の体調などの要因により、測定法毎の結果に差が出る可能性がある。これを防ぐために、1 日に全ての測定法を用いた。そ

の方法は、1セッション(約1時間)の間に、各測定法で2ブロックの閾値測定を行った。つまり、1セッションで6ブロックの測定を行なう。そして、合計4セッションの測定を行った。また、全ての被験者が同じ順序で測定を行うと、実験変数(測定法)と余剰変数(疲労効果などの要因)が交絡する。そのため、全ての組合せ順序で実験を行い、その交絡をキャンセルするようにした。

聴取訓練を受けた被験者群での実験の目的は、各測定法で閾値に差があるか、実験効率に差があるかと、被験者の負担に差があるのかを調べることである。そこで、データが安定するまで3セッション以上の予備実験を行い、この実験の刺激音に十分慣れてもらった。その後、本試行を4セッション行なった。実験の始めには毎回約30試行の練習試行を行ってから、本試行を始めた。なお、練習試行では学習効果を高めるために被験者に正誤のフィードバックを行ったが、本試行ではフィードバックを行わなかった。

5.2.2.2 結果

図5.1に各測定法の被験者毎の閾値の平均値とその標準偏差を示し、図5.2に各測定法の被験者全員の閾値の平均値とその標準偏差を示す。

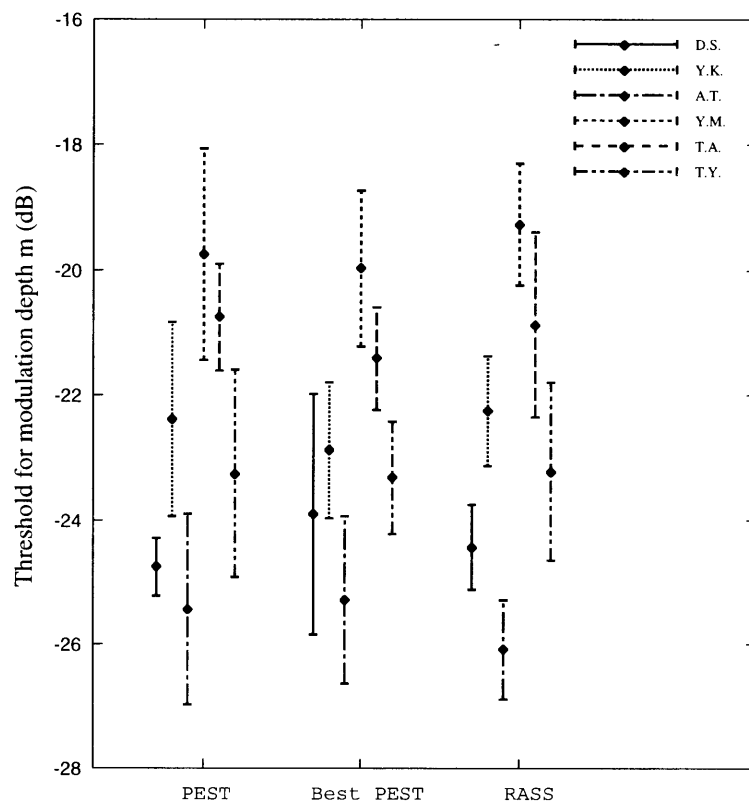


図 5.1: 聴取訓練を受けた被験者 6 人の被験者毎の変調検知閾 (エラーバーは ± 1 標準偏差を示す)

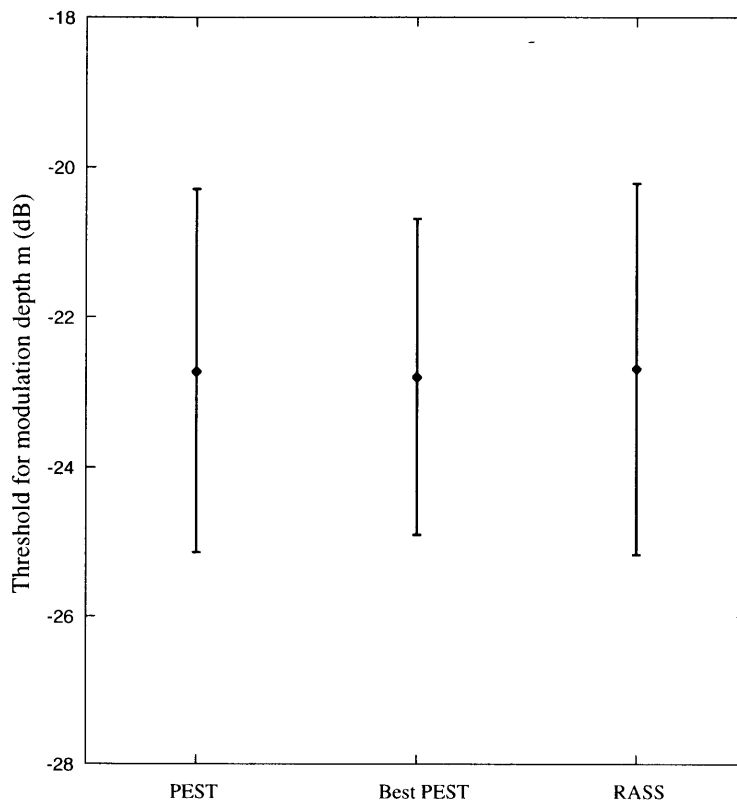


図 5.2: 聴取訓練を受けた被験者の測定法間の平均閾値及び標準偏差 (エラーバーは ± 1 標準偏差を示す)

PEST 法、RASS 法の試行数の平均値及びその標準偏差を表 5.1 に示す。

表 5.1: 聴取訓練を受けた被験者 6 人の PEST 法と RASS 法の試行数の平均値及び、標準偏差

| | Mean trial number | Standard deviation |
|------|-------------------|--------------------|
| PEST | 47.96 | 17.21 |
| RASS | 53.02 | 10.48 |

表 5.1 より、Best PEST 法と試行数を比較すると平均値が 50 試行前後となり、ほぼ同じ実験時間と言える。そのため、各測定法の実験条件はほぼ同じである。また、PEST 法は RASS 法と比較すると試行数の標準偏差は大きな値を取る。そのため、PEST 法は実験終了時間が予測しにくく、実験計画を立てにくいことが分かった。

実験時間がほぼ同じであったため、実験効率は測定で求めた閾値のばらつきで評価できる。そこで、測定法間の閾値について Levene の等分散の検定を行なった。その結果、測定法間の閾値の標準偏差に差はなかった ($F(17,126)=1.58$ 、危険率 5%)。よって、効率が同じである。表 5.2 に分散分析の結果を示す。

表 5.2: 聴取訓練を受けた被験者の閾値の分散分析の結果

| | | | 平方和 | 自由度 | 平均平方 | F 値 | 有意確率 |
|----|---------|-----------|---------|-----|---------|--------|-------|
| 閾値 | 主効果 | 測定法 | 0.281 | 2 | 0.140 | 0.90 | 0.914 |
| | | 被験者 | 566.572 | 5 | 113.314 | 72.432 | 0.00 |
| | 2 次交互作用 | 測定法 × 被験者 | 11.294 | 10 | 1.129 | 0.722 | 0.703 |
| | | 残差 | 197.118 | 126 | 1.564 | | |
| | | 合計 | 775.264 | 143 | 5.421 | | |

また、分散分析の結果、測定法間の閾値の平均値に有意な差がなかった (危険率 5%)。

5.2.2.3 内観報告

被験者の閾値結果のみで、総合的な効率を判断するのではなく、測定法の違いによる刺激音を呈示するアルゴリズムの違いが被験者の負担に及ぼす影響を調べるために内観を報告させた。その方法は、被験者に集中力、疲労感、実験の難しさ、心理的な実験時間について、各ブロックの終了後に、次に述べる 7 段階の評価をしてもらった。また、好みの測定法、実験のやり易さについても評価してもらった。ただし、著者を除いた 5 人の結果である。

- 心理的な実験時間
 - 1: 非常に長い
 - 4: ちょうど良い
 - 7: 非常に短い
- 集中力
 - 1: 最初から集中できなかった
 - 7: 最後まで集中できた
- 疲労感
 - 1: 非常に疲れた
 - 7: 全然疲れなかった

- 実験の難易度

1: 非常に難しかった

4: 普通

7: 非常に簡単

表 5.3に各測定法の内観報告の評価値の平均値を示す。被験者が3種類の適応法に全てについて、それぞれ答えているため、各測定法毎の被験者が感じたことについての相対的な比較ができる。

表 5.3: 聴取訓練を受けた被験者の内観報告の評価値の平均値

| | PEST | BP | RASS |
|----------|------|------|------|
| 心理的な実験時間 | 4.20 | 3.75 | 3.68 |
| 集中力 | 6.15 | 5.88 | 5.88 |
| 実験の難易度 | 4.30 | 3.35 | 4.30 |
| 疲労感 | 5.18 | 4.83 | 5.10 |

ここで、各測定法の被験者群の内観の評価の中央値に差があるかどうかを調べるため、対応のある3条件の中央値の比較をFriedman検定で行なった。その結果、実験の難易度に有意な差があった(危険率1%)。そこで、どの測定法間で有意差があるのかを調べるため、Wilcoxonの符号付き順位検定を実験の難易度で行なった。その結果、「PEST法とRASS法」の間には有意な差がなかった(危険率5%)。そして、「Best PEST法」と「PEST法、RASS法」の間で有意な差(危険率1%)があった。よって、Best PEST法が他の測定法より、被験者が難しく感じていることが分かった。このことは、Best PEST法の刺激呈示のアルゴリズムが他の測定法と根本的に異なるためである。Best PEST法以外の測定法は、徐々に呈示される刺激が閾値に近付いて行くが、逆に、Best PEST法は急激に閾値付近のみを測定するため、被験者は難しいと感じたようである。

PEST法は試行数が少ない場合、非常にやり易かったり、疲労感なども少ないが、反対に試行数が多くなると、非常にやりにくいや非常に疲れたという具合に評価ががらりと変わることが分かった。

さらに、ブロック毎に測定法についての測定法のやり易さを被験者に以下の7段階評価で答えてもらった。

- 1: 非常にやりにくい
- 7: 非常にやり易い

結果を表 5.4 に示す。

表 5.4: 実験のやり易さの評価点の平均値

| | Average |
|-----------|---------|
| PEST | 4.35 |
| Best PEST | 3.65 |
| RASS | 4.58 |

実験のやり易さの評価値の中央値に差があるかを Friedman 検定で調べた。その結果、測定法間の実験のやり易さには、有意な差はなかった (危険率 5%)。

被験者の好みの測定法については、PEST 法と Best PEST 法がそれぞれ 2 名、RASS 法が 1 名だった。よって、被験者毎に好みの測定法は違うことが分かった。

聴取訓練を受けた被験者群の場合、PEST 法は試行数の標準偏差が大きく実験計画を立てにくいいため、あまり最適ではない。Best PEST 法は、他の適応法より難易度が高いことが分かった。以上のことから、比較的实验計画も立てやすく、内観報告からも悪い評価がなかった RASS 法が最適であることが分かった。

5.2.3 聴取訓練を受けていない被験者を用いた実験

5.2.3.1 実験条件

聴取訓練を受けた被験者群の結果より、各測定法の実験効率 (標準偏差及び実験時間) は同じであることが分かった。そこで、共通する実験条件 (5.2.1) を用い、聴取訓練を受けていない被験者群で、適応法の効率比較を行なう。この被験者群は、九州芸術工科大学の音響設計学科以外の学生で、これまでに全く聴取訓練を受けたり、聴取実験に参加したことがない。訓練を受けた被験者群では、各測定法の実験効率を調べるため、被験者全員に全ての測定法で実験を行った。しかし、聴取訓練を受けていない被験者群では、測定法の刺激音を呈示するアルゴリズムの違いによる被験者毎の閾値の安定性を調べる。そのため、同一の被験者が複数の測定法で実験を行なうことによる刺激音の学習の進行を避けるため、同一の被験者は一つの測定法

のみに固定して、測定法の効率を比較する。そこで、第2章のPEST法とBest PEST法の被験者の閾値測定の結果を用いた。そして、新たにRASS法の被験者(19から30歳の健聴者)10人の閾値測定を行ない、適応法の効率及び被験者の負担について検討を行なう。

実験時間は、約1時間とした(この時間区切りを1セッションとする)。1時間で予備試行(正誤のフィードバックあり)を1ブロック測定し、その後、本試行(正誤のフィードバックなし)を3ブロック測定した。また、各ブロックの間には、十分な休憩時間を取った。そして、計3日で3セッションの聴取実験を行い、各測定法で被験者の閾値の安定性などを調べた。ただし、予備試行の結果はデータに取り込まなかった。

5.2.3.2 結果

セッション毎の個人データを図5.3~5.5に示す(1セッションで測定した3ブロックの平均値閾値で示す。ただし、PEST法とBest PEST法の結果は図2.5、図2.6と同じ図である)。また、それぞれの測定法で各10人のセッション間での閾値の平均値と標準偏差を図5.6に示す。

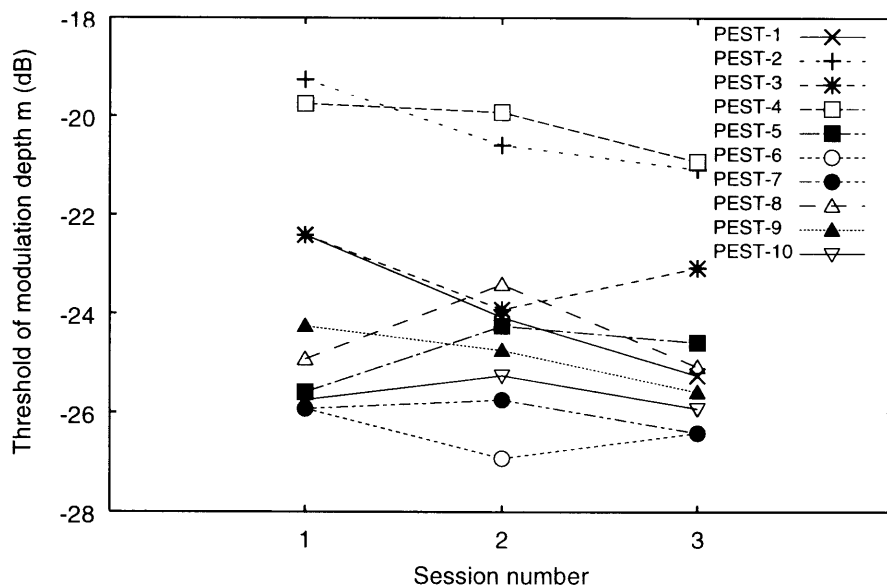


図 5.3: 聴取訓練を受けていない PEST 法の被験者 10 人のセッション間の平均閾値

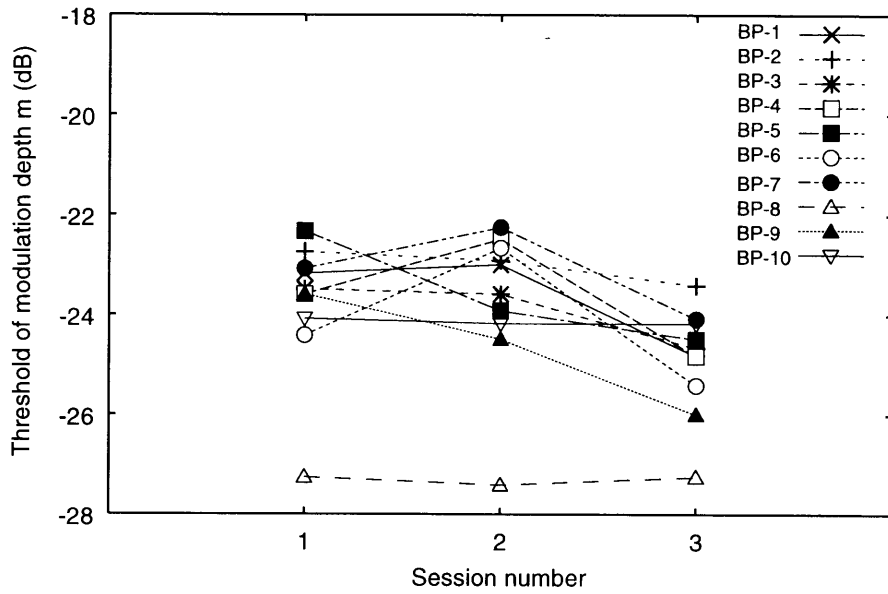


図 5.4: 聴取訓練を受けていない Best PEST 法の被験者 10 人のセッション間の平均閾値

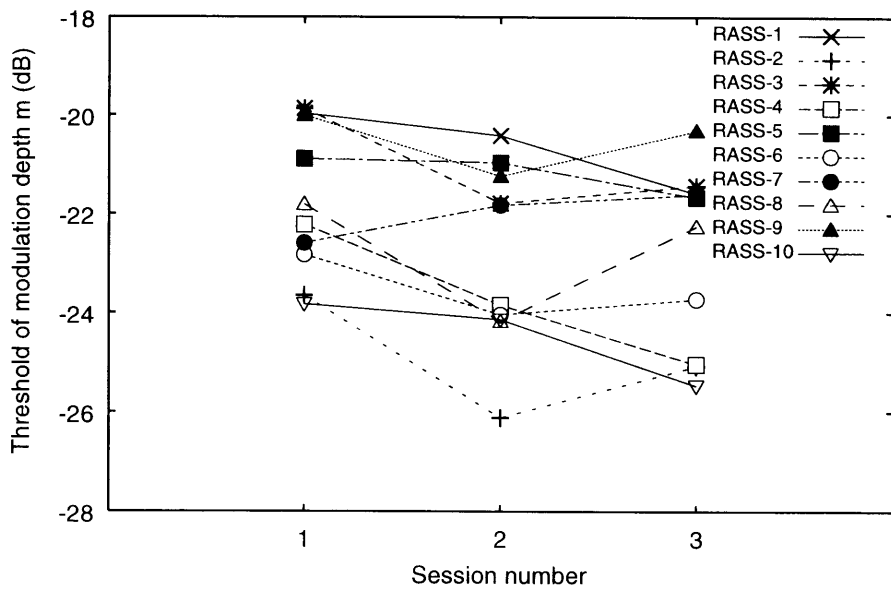


図 5.5: 聴取訓練を受けていない RASS 法の被験者 10 人のセッション間の平均閾値

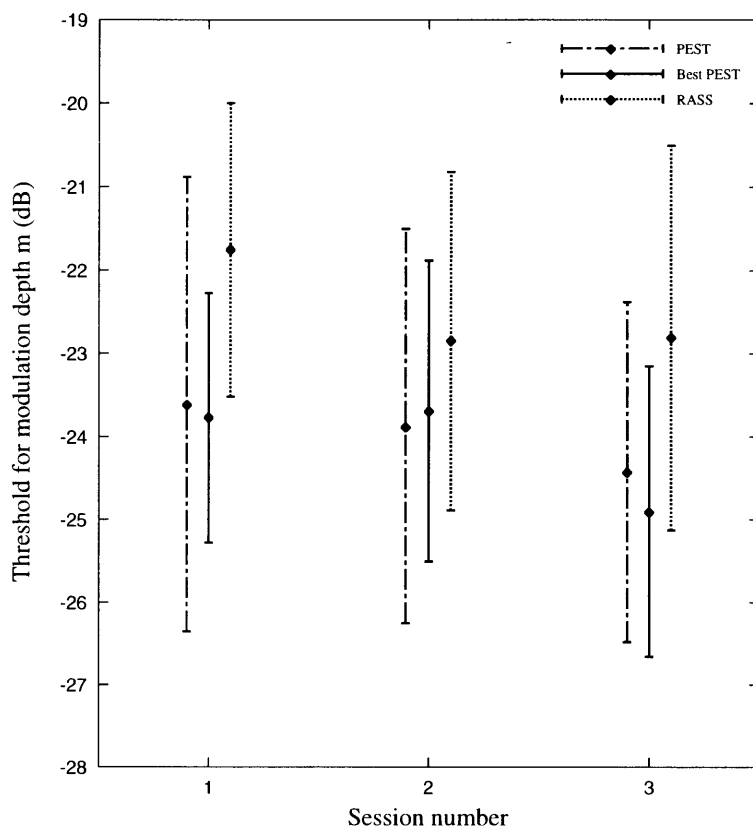


図 5.6: 聴取訓練を受けていない被験者の適応法間のセッション間での平均閾値及び標準偏差 (エラーバーは ± 1 標準偏差を示す)

各測定法の実験効率を調べるために、測定法間の閾値について、セッション間で Levene の等分散の検定を行なった。その結果、有意な差があった ($F(8,261)=2.672$ 、有意差 1%)。

そのため、等分散を仮定しない多重比較を行うことで、どの測定法のセッションと差があるのかを Tamhane T2 検定、Dunnnett T3 検定、Games-Howell 検定を行った。それらの結果、「(PEST 法の 2、3 セッション、Best PEST 法の 1、2、3 セッション) と (RASS の 1 セッション)」、「(Best PEST 法の 3 セッションと RASS 法の 1、2、3 セッション)」で閾値に有意な差 (1%) があった。

これは、訓練を受けた被験者群の結果より、被験者個人の閾値に有意な差があったことから考察できる (表 5.2)。測定法毎で、聴取訓練を受けていない被験者群の閾値に隔たりがあったため、等分散の検定に通らなかったと推察できる。そこで、各被験者個人の 1 セッション目の閾値を基準にし、2、3 セッション目で閾値がどのように変化したのかを調べる (図 5.7~5.10、ただし、PEST 法と Best PEST 法の結果は、

図 2.9、図 2.10と同じである)。

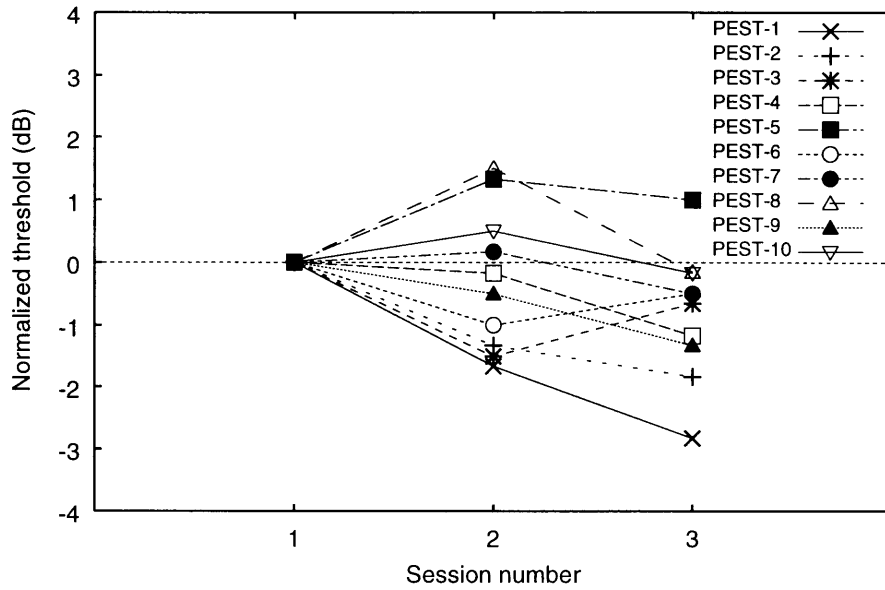


図 5.7: 1セッション目を基準にした PEST 法の各被験者の閾値の推移

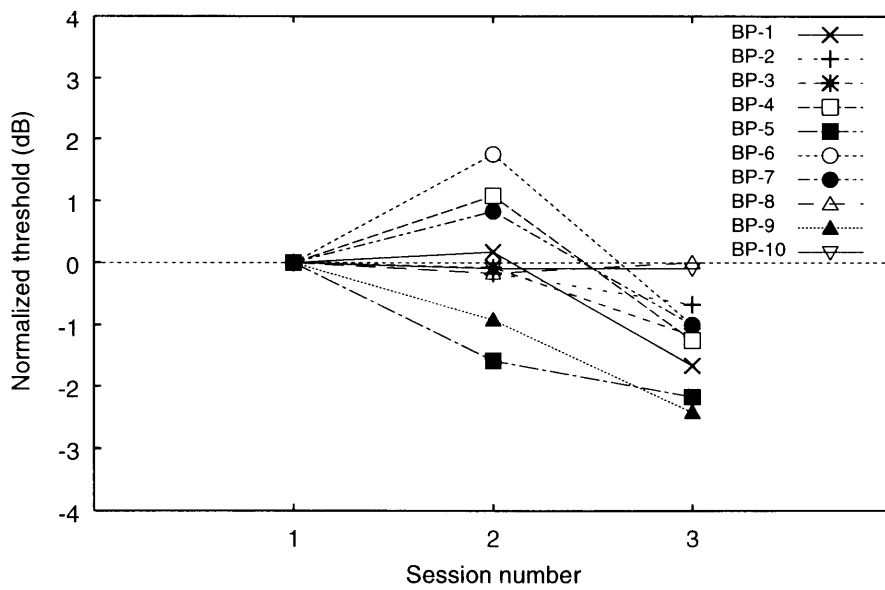


図 5.8: 1セッション目を基準にした Best PEST 法の各被験者の閾値の推移

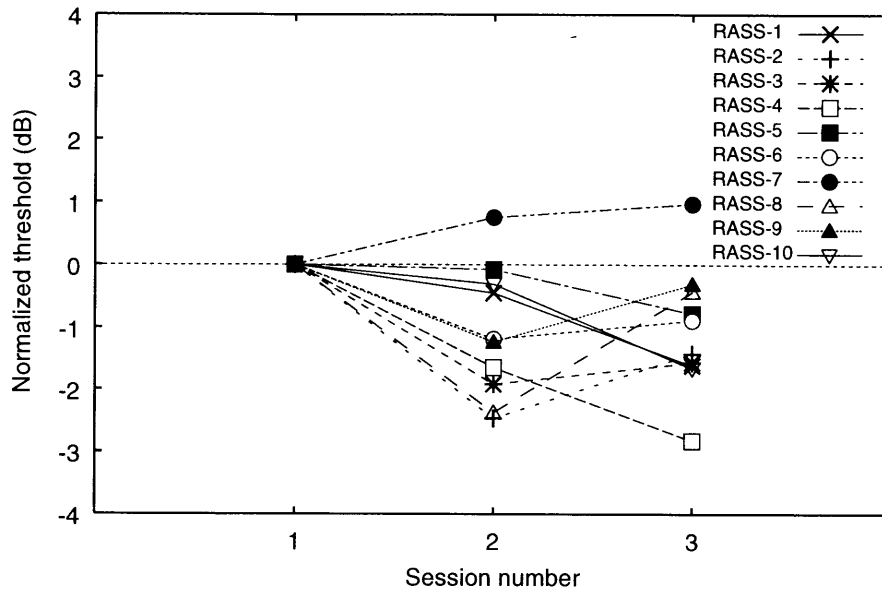


図 5.9: 1セッション目を基準にした RASS 法の各被験者の閾値の推移

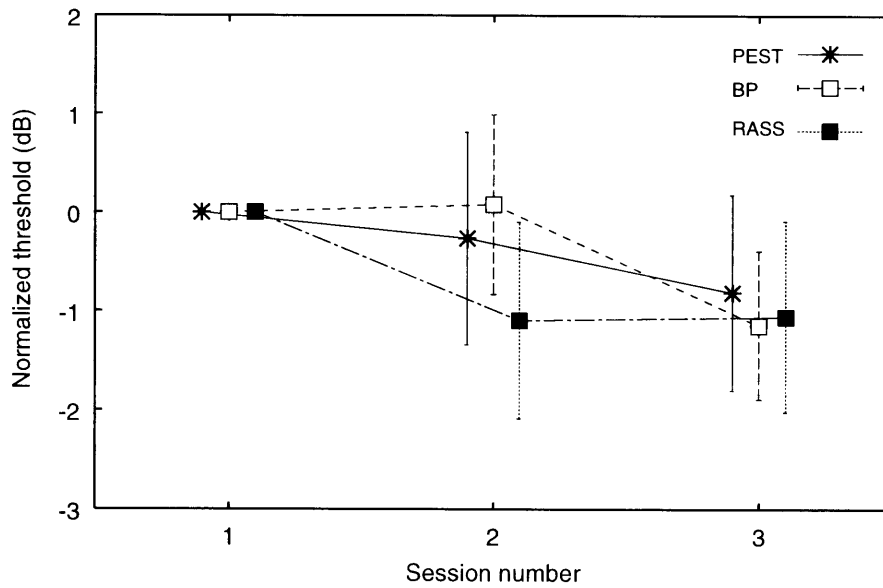


図 5.10: 1セッション目を基準にした各適応法の平均閾値と標準偏差の推移 (エラーバーは ± 1 標準偏差を示す)

そこで、Levene の等分散の検定を被験者毎に 1セッション目の閾値で基準化した 2、3セッション目で別々に行なった。その結果、2セッション目は測定法毎の閾値の分散に有意な差はなかった ($F(2,27)=0.667$ 、危険率 5%)。また、3セッション目

も測定法毎の閾値の分散に有意な差はなかった ($F(2,27)=0.314$ 、危険率5%)。そのため、効率は同じである。今回は被験者数が若干少なかったため、測定法間の被験者で閾値が隔たる被験者が集まり、等分散の検定で閾値に有意な差があった。しかし、1セッション目からの閾値の推移では、分散に差がないことが分かった。

表 5.5: 聴取訓練を受けていない被験者群の PEST 法と RASS 法の平均試行数及び標準偏差

| | Mean trial number | Standard deviation |
|--------------------|-------------------|--------------------|
| PEST (1st session) | 57.4 | 12.5 |
| PEST (2nd session) | 52.5 | 12.5 |
| PEST (3rd session) | 53.8 | 11.2 |
| RASS (1st session) | 55.4 | 10.2 |
| RASS (2nd session) | 54.9 | 11.2 |
| RASS (3rd session) | 56.5 | 10.0 |

PEST 法の平均試行数は、他の測定法 (50 試行) と比較しても、ほぼ同じである。しかし、PEST 法は測定中に被験者が標準刺激と比較刺激の差が分からなくなり、測定を中断した割合が約 2 割となり、他の測定法と比較すると測定時間が長くかかった。また、RASS 法の試行数の標準偏差は訓練を受けた被験者群とあまり変わらなかった (表 5.1、5.5)。これより、RASS 法は被験者の違いによらず、試行数の標準偏差に変わりがないことが分かった。

5.2.3.3 内観報告

被験者の閾値結果のみで、総合的な効率を判断するのではなく、測定法の違いによる刺激音の呈示するアルゴリズムの違いが、被験者の負担に及ぼす影響を調べるために内観を報告してもらった。その方法は、被験者に集中力、疲労感、実験の難しさ、心理的な実験時間について、各ブロックまたはセッションの終了後に、次に述べる 7 段階の評価をしてもらった。

- 心理的な実験時間 (1 ブロック毎に集計)

1: 非常に長い

- 4: ちょうど良い
- 7: 非常に短い
- 集中力 (1 ブロック毎に集計)
 - 1: 最初から集中できなかった
 - 7: 最後まで集中できた
- 疲労感 (1 ブロック毎に集計)
 - 1: 非常に疲れた
 - 7: 全然疲れなかった
- 実験の難易度 (1 セッション毎に集計)
 - 1: 非常に難しかった
 - 4: 普通
 - 7: 非常に簡単

表 5.6: 聴取訓練を受けていない被験者の適応法の内観報告の各評価の平均値

| | PEST | Best PEST | RASS |
|----------|------|-----------|------|
| 心理的な実験時間 | 4.00 | 4.30 | 4.46 |
| 集中力 | 5.00 | 4.81 | 5.90 |
| 疲労感 | 5.21 | 4.63 | 5.73 |
| 実験の難易度 | 3.63 | 3.50 | 3.80 |

表 5.6に各測定法の内観報告の平均値を示す。各測定法の評価値の中央値に差があるかどうかを調べるために、Kruskal Wallis 検定を行なった。その結果、心理的な実験時間、集中力、疲労感で各適応法間で有意な差があった(危険率1%)。そこで、どの測定法間で評価値の中央値に有意な差があったのかを調べるために、Mann-Whitney 検定を行なった。その結果、心理的な時間は、「PEST 法と Best PEST 法」の間には有意な差がなかった(危険率5%)。しかし、「PEST 法と Best PEST 法」と「RASS 法」の間に有意な差があった(危険率1%)。よって、RASS 法が他の測定法より、心理的な実験時間を短く感じていることが分かった。集中力は、「PEST 法と Best PEST 法」の間に有意な差はなかった(危険率5%)。しかし、「PEST 法と Best PEST 法」と「RASS 法」の間に有意な差があった(危険率1%)。よって、RASS が他の測定法

より、集中力が持続することが分かった。疲労感は全ての測定法間でそれぞれ有意な差があった(危険率1%)。よって、RASS法が最も疲労感を少なく感じ、次いで、PEST法、Best PEST法の順となった。

この内観報告結果より、3種類の適応法の中ではRASS法が最も評価が良かった。PEST法とBest PEST法では、第2章と同様にBest PEST法が疲労感にのみ、悪い評価が得られた。聴取訓練を受けていない被験者群を対象とした場合、RASS法が最も被験者の負担が少なく、適していることが分かった。

5.2.4 聴取訓練を受けた被験者を用いた心理測定関数の傾き推定誤差実験

5.2.4.1 実験条件

Best PEST法とRASS法は心理測定関数の傾き推定値を使用する。そのため、傾き推定値の推定誤差によって、効率が変化するかを実験で調べる必要がある。シミュレーションでは、RASS法は傾き推定値が真値の1/2の時に標準偏差が大きく、若干推定値が大きくても変化があまりなかった。Best PEST法は若干傾き推定値が小さい時に、バイアスが小さかった。そこで、シミュレーションで得られたバイアス及び標準偏差の差が実験で出るかを調べる。

被験者は聴取訓練を受けた3名(21、25、28歳)であり、まず、心理測定関数の傾き値を恒常法で測定した。刺激音はこれまでの実験と同様に振幅変調検知閾の実験である。刺激レンジを12dBとし、7刺激レベルで280試行を用い、最尤推定法で心理測定関数(ロジスティック関数)を当てはめ、傾き値と閾値を求めた。結果を表5.7に示す。

表 5.7: 恒常法で測定した被験者3名の心理測定関数の傾き値及び閾値

| subject | slope | threshold (dB) |
|---------|-------|----------------|
| AT | 1.2 | -25.58 |
| TA | 0.55 | -21.56 |
| YK | 0.45 | -22.85 |

この心理測定関数の傾き値を基準(真値)にし、それぞれの被験者毎にRASS法、Best PEST法で使用する傾き値を真値、真値の1/2倍(underestimate)、真値の2倍(overestimate)の計6種類の測定条件を用いて実験を行なった。そのため、被験者毎

に RASS 法では刺激レベルの変化幅が異なる。なお、実験条件は共通する実験条件 (5.2.1) と同様である。1セッションで全ての組合せ (6通り) を使用し、計6セッションの測定を行なった。

5.2.4.2 結果

結果を図 5.11、5.12に示す。Levene の等分散の検定を行なったが閾値に有意な差はなかった ($F(17,90)=1.445$ 、危険率 5%)。その後、分散分析を行なったが6種類の条件で測定法間の閾値に有意な差 (危険率 5%) はなかった (表 5.8)。

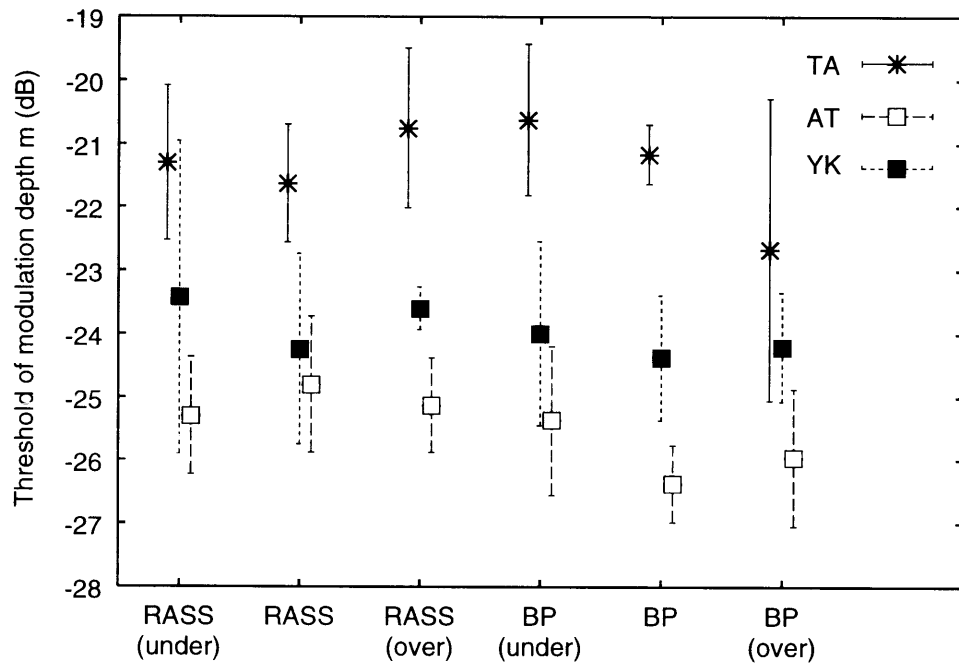


図 5.11: 心理測定関数の傾き推定値が mismatching の場合の各被験者毎の平均閾値及び標準偏差 (エラーバーは ± 1 標準偏差である。心理測定関数の傾き推定値は真値、1/2 倍 (under)、2 倍 (over) を使用した場合の結果を示す。)

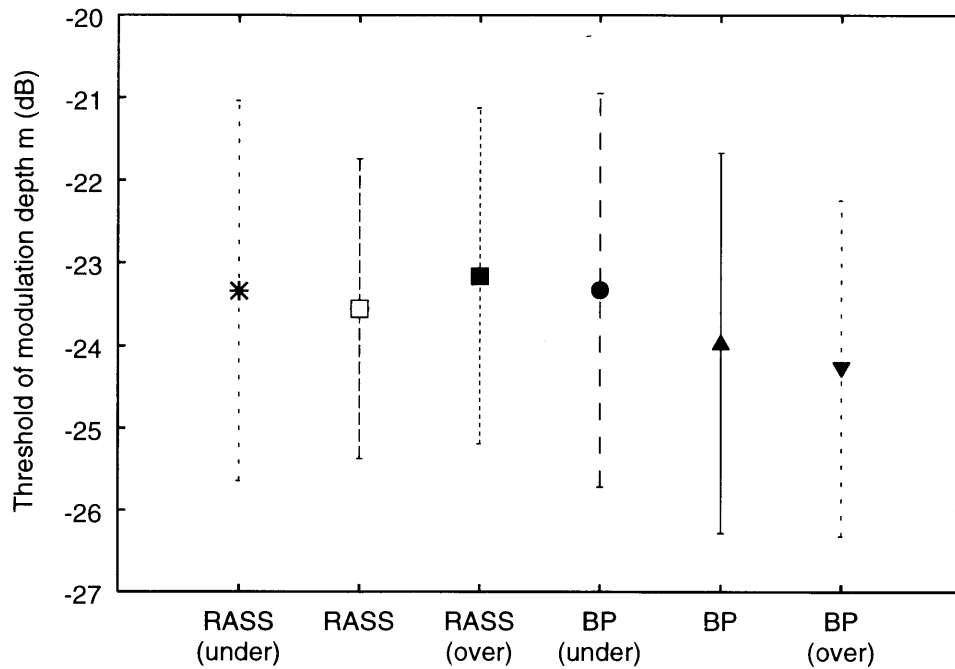


図 5.12: 心理測定関数の傾き推定値がミスマッチングの場合の各測定法の閾値平均値 (エラーバーは ± 1 標準偏差である。心理測定関数の傾き推定値は真値、1/2 倍 (under)、2 倍 (over) を使用した場合の結果を示す。)

表 5.8: 心理測定関数の傾き推定値が閾値に及ぼす影響についての分散分析の結果

| | | 平方和 | 自由度 | 平均平方 | F 値 | 有意確率 |
|---------|------------------|---------|-----|---------|--------|-------|
| 閾値 | 測定法 | 16.673 | 5 | 3.335 | 2.060 | 0.078 |
| | 被験者 | 314.718 | 2 | 157.359 | 97.197 | 0.00 |
| 2 次交互作用 | 測定法 \times 被験者 | 13.956 | 10 | 1.396 | 0.862 | 0.571 |
| | 残差 | 145.708 | 90 | 1.619 | | |
| | 合計 | 491.055 | 107 | 4.589 | | |

表 5.9: 心理測定関数の傾き推定値がミスマッチングの場合、RASS 法の 1 ブロックに要する平均試行数及び標準偏差

| | Mean trial number | Standard deviation |
|----------------------|-------------------|--------------------|
| RASS (underestimate) | 43.7 | 10.6 |
| RASS (normal) | 54.3 | 12.8 |
| RASS (overestimate) | 76.5 | 22.5 |

RASS の傾き推定値が真値の 2 倍の場合、試行数の平均値及び標準偏差が大きくなった。この理由として、被験者 AT の心理測定関数の傾き値の真値 ($a=1.2$) が大きかったためである。100% 正応答率が得られた場合のステップ幅が 0.75(dB) となり、約 6dB 閾値上から測定を開始したため、閾値付近までに到達するまで、最低 8 段階の刺激レベルの変化を要する。そのため、閾値付近まで到達する最低試行数が約 40 試行 (8 段階 \times 5 試行) 必要だったためである。RASS 法の各正応答率のステップ幅はあらかじめ計算できるため、閾値付近の刺激レベルまでの最低試行数が極端に多くなる場合、測定を開始する刺激レベルを下げれば、試行数が増加する問題は解消される。

心理測定関数の傾き推定値は真値の 1/2 から 2 倍の範囲で閾値に有意な差はなかったが、RASS 法は傾き推定値が真値の 2 倍の場合、試行数が増加することが分かった。

5.3 まとめ

聴取訓練を受けた被験者と受けていない被験者群で変調検知閾値を 3 種類の適応法 (PEST 法、Best PEST 法、RASS 法) で測定した。その結果、以下のことが分かった。

- 聴取訓練を受けた被験者群では、測定法の違いによる閾値の差はなく、実験効率も同じあることが分かった。内観報告より、被験者毎に好みの測定法が違うことが分かった。Best PEST 法は閾値付近の刺激レベルのみを呈示するため、他の測定法より難しく感じていることが分かった。PEST 法は測定で要する試行数によって、疲労感ややり易さの評価が極端に変わることが分かった。PEST 法は、試行数の標準偏差も大きく、測定に要する実験時間が予測できないため、

実験計画を立てにくいことが分かった。以上のことから、比較的实验計画も立てやすく、内観報告からも悪い評価がなかった RASS 法が最適であることが分かった。

- 聴取訓練を受けていない被験者群では、測定法毎の被験者個人の閾値にばらつきがあったため、等分散の検定で閾値に有意な差があったが、1セッション目を基準とした、2、3セッション目での閾値の分散は同じであることが分かった。内観報告の結果から、最も評価が良かったのは RASS 法である。Best PEST 法は実験の疲労感で最も悪い評価であった。これは、閾値付近の刺激レベルのみを呈示するためであると考えられる。PEST 法は測定の終了が予測できず、実験を取り直した割合も多く、測定時間が他の測定法より長くかかった。よって、総合的な評価をすると RASS 法が聴取訓練を受けていない被験者群にも適していることが分かった。
- Best PEST 法及び RASS 法は、心理測定関数の傾き推定値が真値の $1/2$ から 2 倍の範囲でも閾値に有意な差がないことが分かった。しかし、RASS 法は傾き推定値が 2 倍の場合、試行数が増加することが分かった。

第 6 章

まとめ

本研究では、聴覚心理実験で 2IFC 法 (2 者強制選択法) を使用し、効率良く閾値を測定するため、実験効率と被験者の負担の点でいくつかの心理測定法で検討を行い、新しい適応法 (RASS 法) を提案した。

第 1 章では、本研究の背景及び目的を述べた。

第 2 章では、従来からある恒常法と適応法 (PEST 法、Best PEST 法) の効率を変調検知閾実験で比較した。聴取訓練を受けた被験者群では、測定法の違いによる閾値の差はなく、実験効率も同じであることが分かった。そして、被験者毎に好みの測定法が違うことが分かった。また、聴取訓練を受けていない被験者群では、測定法の違いによる閾値の差はなく、実験効率は同じであることが分かった。しかし、恒常法は適応法と比較すると、被験者毎の閾値がセッション間で大きく変動する被験者が多く、効率が悪いことが分かった。また、PEST 法は、測定時間が他の測定法よりも長くなり、効率が悪い。Best PEST 法の効率は良かったが、一方、内観報告では、PEST 法の被験者が他の測定法の被験者より、疲労感を少なく感じていることが分かった。

第 3 章では、Best PEST 法より効率が良く、PEST 法より被験者の負担が少ない RASS 法 (Rapid Adjustment of Step Size Method) を新たに開発し、その理論について述べた。また、RASS 法は、刺激レベルの変化方法のアルゴリズムが非常に単純であり、測定の処理時間が速い。刺激レベルの変化がチャンスレベル付近では、起こりやすく、比較的分かりやすい刺激レベルが出やすくなっている。また、心理測定関数をロジスティック関数と仮定し、その傾き推定値を使用することで、測定条件のパラメータが簡単に決定できるという特徴がある。

第 4 章では、モンテ・カルロ・シミュレーションで適応法 (PEST 法、Best PEST 法、RASS 法) の効率を調べた。刺激レンジが狭い場合には、Best PEST 法の効率

が良いが、刺激レンジが広がるにつれ、RASS法の効率が良くなることが分かった。このことは、実験を行なう上で、閾値上の刺激レベルなら、どの刺激レベルから測定を開始しても良いという利点があり、予備実験で被験者の閾値を測定する場合、明らかに分かる刺激レベルから測定を開始しても、試行数が若干多くなるだけで、バイアスと標準偏差の値は変わらない。これに対し、Best PEST法は刺激レンジが広がるにつれ、バイアスと標準偏差は大きな値を取り、効率が悪くなるため、刺激レンジを広く設定することができない。また、Best PEST法は、被験者の閾値が設定した刺激レンジ内に入っていない場合、その試行が無駄になるという欠点もある。PEST法は、測定に要する平均試行数の標準偏差が大きな値を取るが、RASS法は、PEST法の約半分以下の標準偏差値を取り、測定時間が予測しやすいため、実験計画を立てやすいことが分かった。そして、さまざまな条件の時にRASS法が、最も効率が良いことが分かった。

第5章では、第2章と同じ変調検知閾実験を行ない、3種類の適応法(PEST法、Best PEST法、RASS法)で効率を比較した。聴取訓練を受けた被験者群では、測定法の違いによる閾値の差はなく、実験効率も同じあることが分かった。内観より、被験者毎に好みの測定法が違ってくることが分かった。そして、Best PEST法が、他の測定法より難しく感じていることが分かった。PEST法は、試行数の標準偏差も大きく、測定に要する実験時間が予測できないため、実験計画を立てにくいことが分かった。以上のことから、比較的实验計画も立てやすく、内観報告からも悪い評価がなかったRASS法が最適であることが分かった。聴取訓練を受けていない被験者群では、測定法毎の被験者個人の閾値にばらつきがあったため、等分散の検定で閾値に有意な差があったが、1セッション目を基準とした、2、3セッション目での閾値の分散は同じであることが分かった。内観報告の結果から、最も評価が良かったのはRASS法である。そのため、総合的な評価をするとRASS法が聴取訓練を受けていない被験者にも適していることが分かった。Best PEST法及びRASS法は、心理測定関数の傾き値を測定で使用する。その心理測定関数の傾き推定値が真値の1/2倍、2倍の場合でも閾値に差がないことが分かった。

本研究で提案したRASS法の特徴は以下の通りである。

- 2IFC法で効率が良い。
- 刺激レベルの変化方法のアルゴリズムが非常に単純である。実験前にステップ幅を決定するため、測定に要する処理が速い。

- 心理測定関数をロジスティック関数と仮定し、その傾き推定値を使用することで、測定条件のパラメータが簡単に決定できる。
- 刺激レベルの変化方向が逆転する山・谷の折り返し数が偶数の場合に、効率が良い(バイアスが小さい)。
- 被験者が応答の押し間違いをしても、同じ刺激レベルで1回までなら、次に呈示される刺激レベルが下がるため、ある程度 lapse に対応できる。
- 刺激レベルの変化が起こりやすく、また、チャンスレベル付近の刺激レベルでは、すぐに分かりやすい刺激レベルが呈示される(50%以下の確率事象が出る確率が高く、その場合、ステップ幅も大きい)ため)。
- 測定試行数の標準偏差が小さく、測定に要する時間が予測しやすいため、実験計画が立てやすい。
- 被験者の負担が少ない。
- シミュレーションの場合、刺激レンジが広い場合、及びさまざまな条件の場合に最も効率が良い。
- 聴取実験の場合、聴取訓練を受けた被験者と受けていない被験者群で、最も最適である。

謝辞

本研究を進めるにあたり、熱心にご指導頂いた九州芸術工科大学 津村 尚志教授に深く感謝いたします。また、本論文をまとめるにあたり、多くの貴重なご助言を頂いた九州芸術工科大学 岩宮 眞一郎教授、山下 由己男助教授に深く感謝いたします。

本研究の RASS 法に関する有益なご助言を頂いた、九州芸術工科大学 角 俊雄助教授、中島 祥好助教授、大阪芸術大学 山田 真司助手に深く感謝いたします。

第2章の実験を手伝って頂いた三芳 弘子氏に深く感謝します。また、実験を進めるにあたり、お忙しい中、被験者をして頂いた皆さんに感謝します。特に第5章の被験者をして頂いた九州芸術工科大学 津村研究室の赤木 太郎氏、津村研究室の卒業生の吉住 剛氏、中島研究室の金深 京子氏、小森 由宗氏、末富 大剛氏、村北 泰規氏に感謝いたします。

参考文献

- [1] J. P. ギルホード著, “精神測定法,” 培風館, 昭和 34 年.
- [2] “新版 心理学辞典,” 平凡社, 1981 年.
- [3] “聴覚心理学概論,” B. C. J. ムーア, 誠信書房, 1994 年.
- [4] 津村 尚志, “最近の聴覚心理実験における新しい測定法,” 音響学会誌 40, 45-51 (1984).
- [5] H. Levitt, “Transformed Up-Down Methods in Psychoacoustics,” J. Acoust. Soc. Am. 49, 467-477 (1970).
- [6] M. M. Taylor and C. Douglas Creelman, “PEST : Efficient Estimates on Probability Functions,” J. Acoust. Soc. Am. 41, 782-787, (1967).
- [7] J. L. Hall, “Hybrid adaptive procedure for estimation of psychometric functions,” J. Acoust. Soc. Am. 69, 1763-1769 (1981).
- [8] Alex P. Pentland, “Maximum likelihood estimation : The best PEST,” Percept. Psychophys. 28, 377-379 (1980).
- [9] William A. Simpson, “The step method: A new adaptive psychophysical procedure,” Percept. Psychophys. 45, 572-576 (1989).
- [10] C. Karenbach, “A single-interval adjustment-matrix (SIAM) procedure for unbiased adapting testing,” J. Acoust. Soc. Am. 88, 2645-2655 (1990).
- [11] 寺岡 章人, 津村 尚志, “訓練を受けた聴取者と受けていない聴取者を用いた三つの心理物理測定法の実験効率の比較,” 音響学会誌 54, 127-133 (1998).
- [12] William A. Simpson, “The method of constant stimuli is efficient,” Percept. Psychophys. 44, 433-436 (1988).

- [13] Andrew B. Watson and Andrew Fitzhugh, "The method of constant stimuli is inefficient," *Percept. Psychophys.* 47, 87-91 (1990).
- [14] A. Hesse, "Comparison of Several Psychophysical Procedures with Respect to Threshold Estimates, Reproducibility and Efficiency," *Acustica* 59, 263-273 (1986).
- [15] J. A. Stillman, "A comparison of three adaptive psychophysical procedures using inexperienced listeners," *Percept. Psychophys.* 46, 345-350 (1989).
- [16] Neal F. Viemeister, "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* 66, 1364-1380 (1979).
- [17] 寺西 立年, "九州芸術工科大学における音響心理学の教育," *音響学会誌* 43, 890-893 (1987).
- [18] Kitamura O. , Sasaki M. , Tsumura. M. , Arita K. , "TRAINING OF HEARING," *INTERNATIONAL CONGRESS ON ACOUSTICS MADRID 4/9-VII*, 336 (1977).
- [19] 北村 音一, 佐々木 貫, 津村 尚志, "調能形成のための訓練について (2) —音圧レベル、周波数特性の判定訓練効果—," *音講論集*, 3-1-4, (1973.5).
- [20] C. Kaernbach, "Simple adaptive testing with the weighted up-down method," *Percept. Psychophys.* 49, 227-229 (1991).
- [21] J. K. O'regan and R. Humbert, "Estimating psychometric functions in forced-choice situations: Significant biases found in threshold and slope estimations when small samples are used," *Percept. Psychophys.* 46, 434-442 (1989).
- [22] William H. Press ら著, 丹慶勝市ら訳, "NUMERICAL RECIPES in C[日本語版]," 技術評論社, 平成 6 年.

付録 A

第 3 章のシミュレーション結果

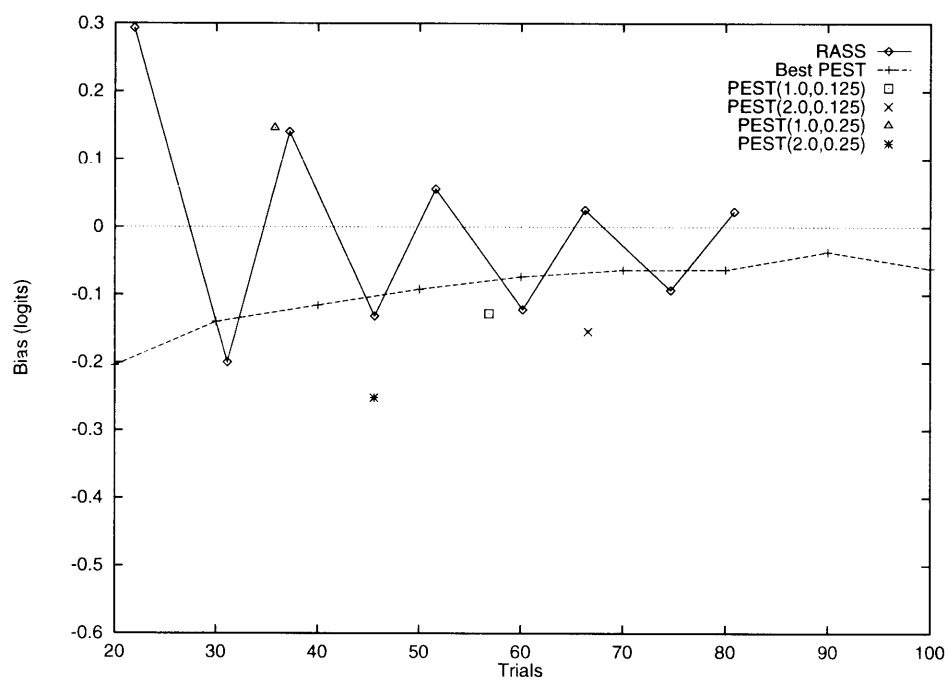


図 A.1: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 5(logits) である。)

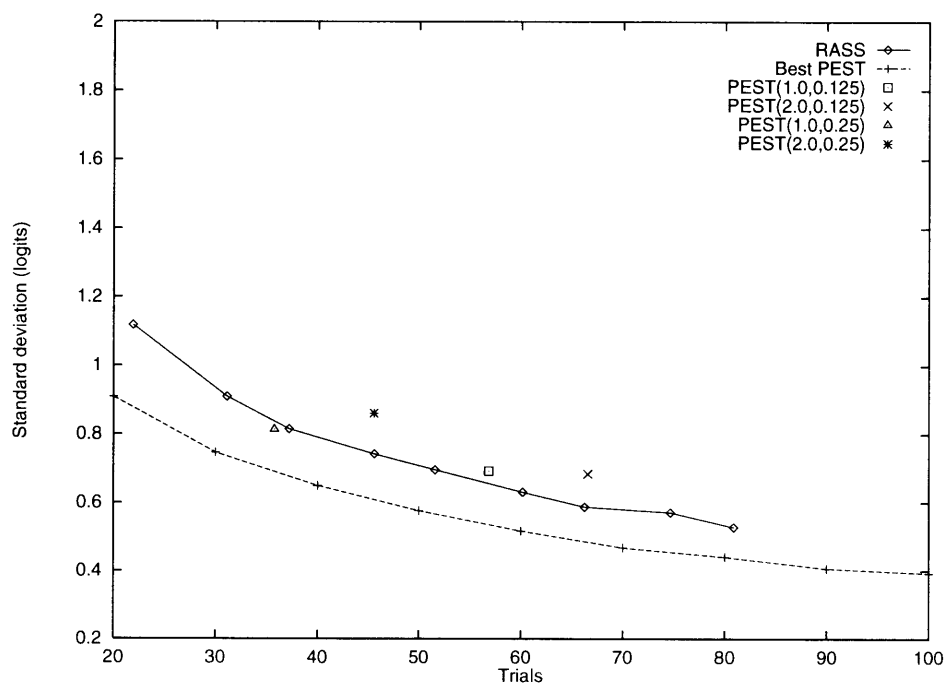


図 A.2: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 5(logits) である。)

表 A.1: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 5(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 21.9 | 31.1 | 37.2 | 45.6 | 51.6 | 60.2 | 66.3 | 74.7 | 80.9 |
| Standard deviation | 7.2 | 8.6 | 9.0 | 10.6 | 11.0 | 12.0 | 12.5 | 13.8 | 13.8 |

表 A.2: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 5(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 56.8 | 27.2 |
| PEST(2.0,0.125) | 66.6 | 34.3 |
| PEST(1.0,0.25) | 35.7 | 19.5 |
| PEST(2.0,0.25) | 45.6 | 23.1 |

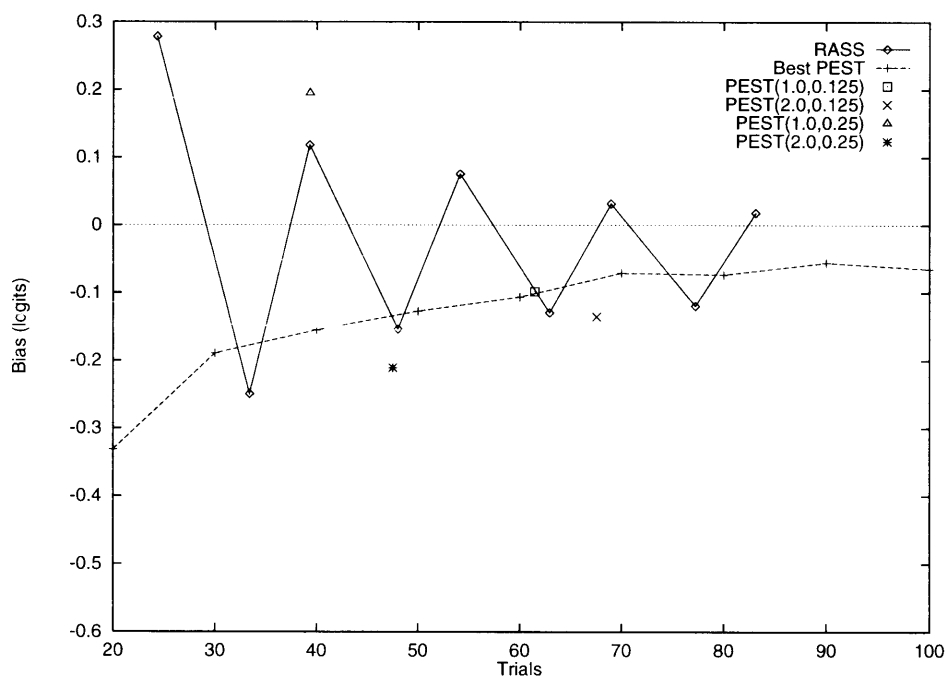


図 A.3: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 7(logits) である。)

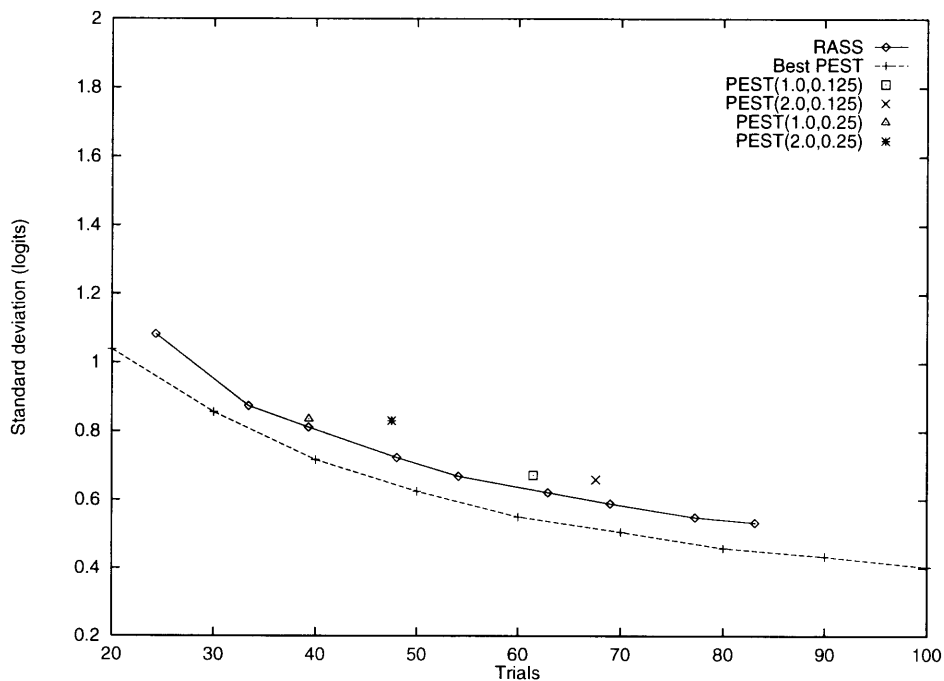


図 A.4: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 7(logits) である。)

表 A.3: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 7(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 24.3 | 33.4 | 39.3 | 48.0 | 54.1 | 62.9 | 69.0 | 77.2 | 83.1 |
| Standard deviation | 7.0 | 8.5 | 9.4 | 10.5 | 11.1 | 12.4 | 12.6 | 13.6 | 13.9 |

表 A.4: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 7(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 61.5 | 27.4 |
| PEST(2.0,0.125) | 67.6 | 34.9 |
| PEST(1.0,0.25) | 39.3 | 19.9 |
| PEST(2.0,0.25) | 47.5 | 22.9 |

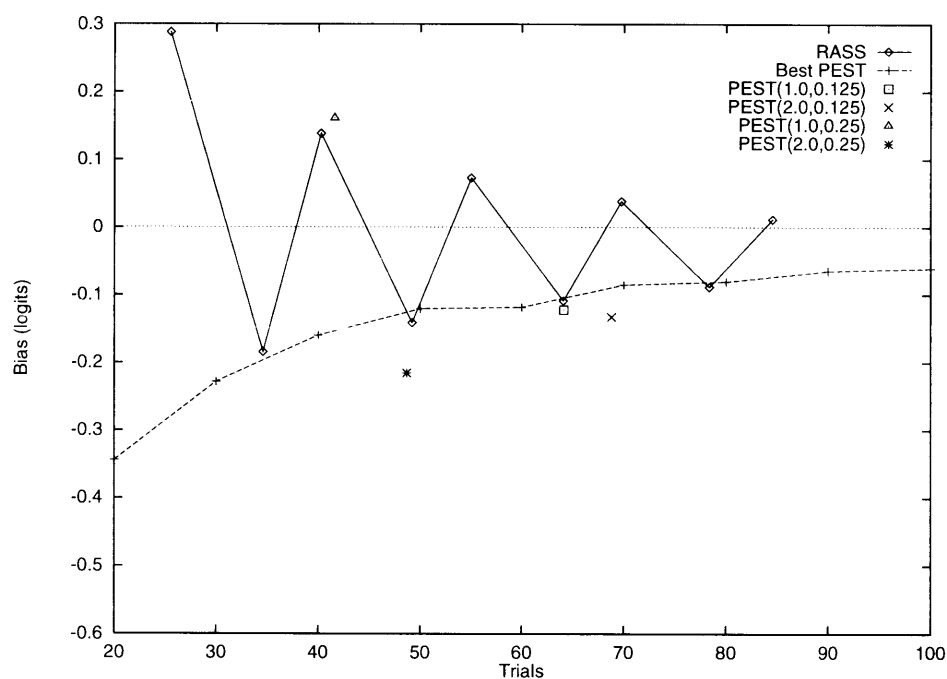


図 A.5: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 8(logits) である。)

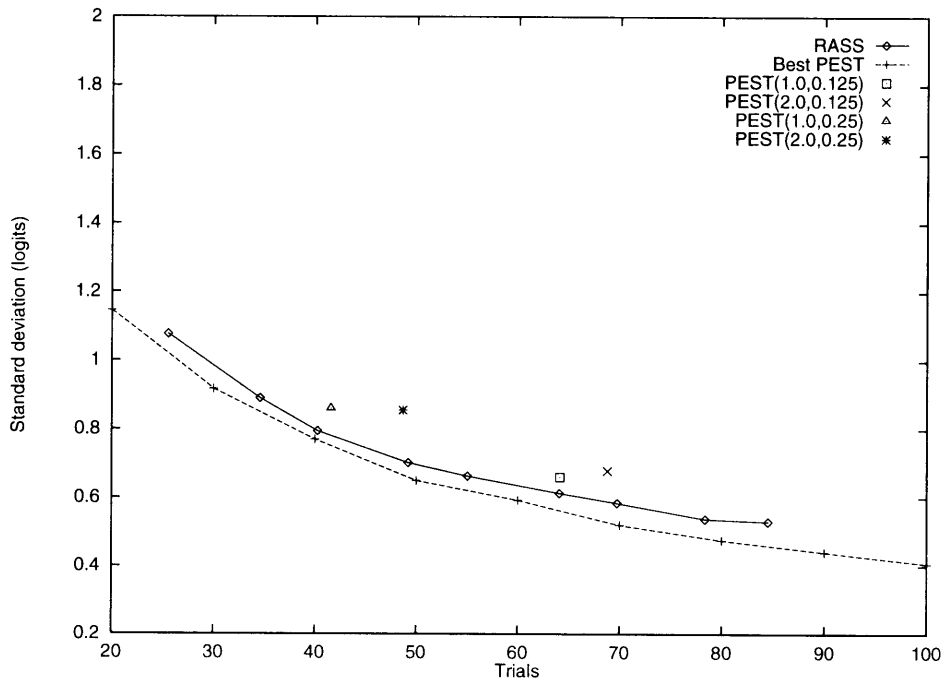


図 A.6: 3種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 8(logits) である。)

表 A.5: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 8(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 25.5 | 34.6 | 40.3 | 49.2 | 55.0 | 64.1 | 69.7 | 78.4 | 84.5 |
| Standard deviation | 7.1 | 8.7 | 9.0 | 10.6 | 10.9 | 12.5 | 12.7 | 13.7 | 14.0 |

表 A.6: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 8(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 64.1 | 28.1 |
| PEST(2.0,0.125) | 68.8 | 34.3 |
| PEST(1.0,0.25) | 41.6 | 19.2 |
| PEST(2.0,0.25) | 48.7 | 22.6 |

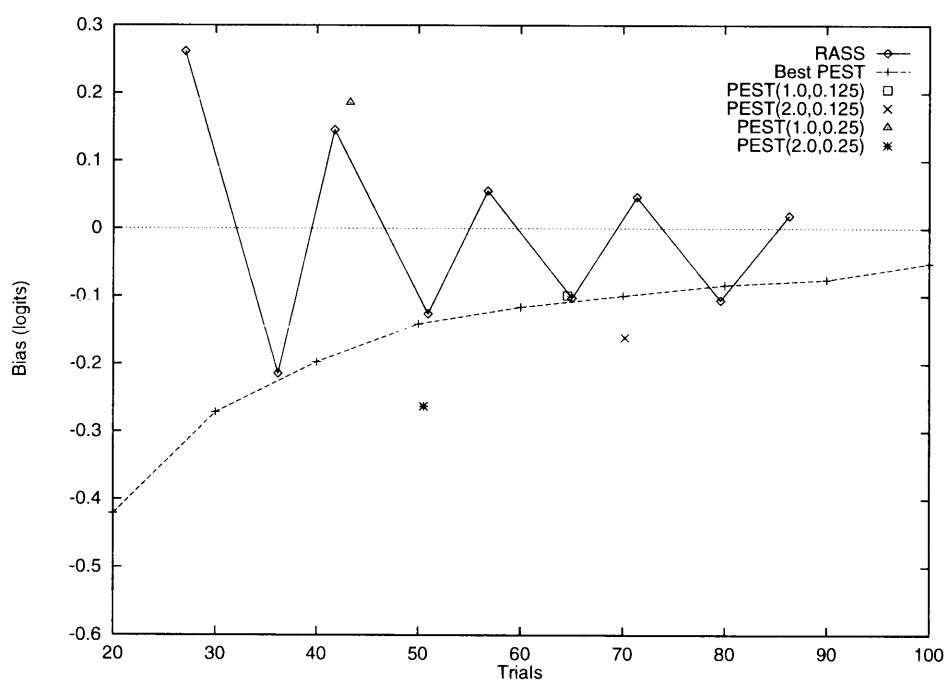


図 A.7: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 9(logits) である。)

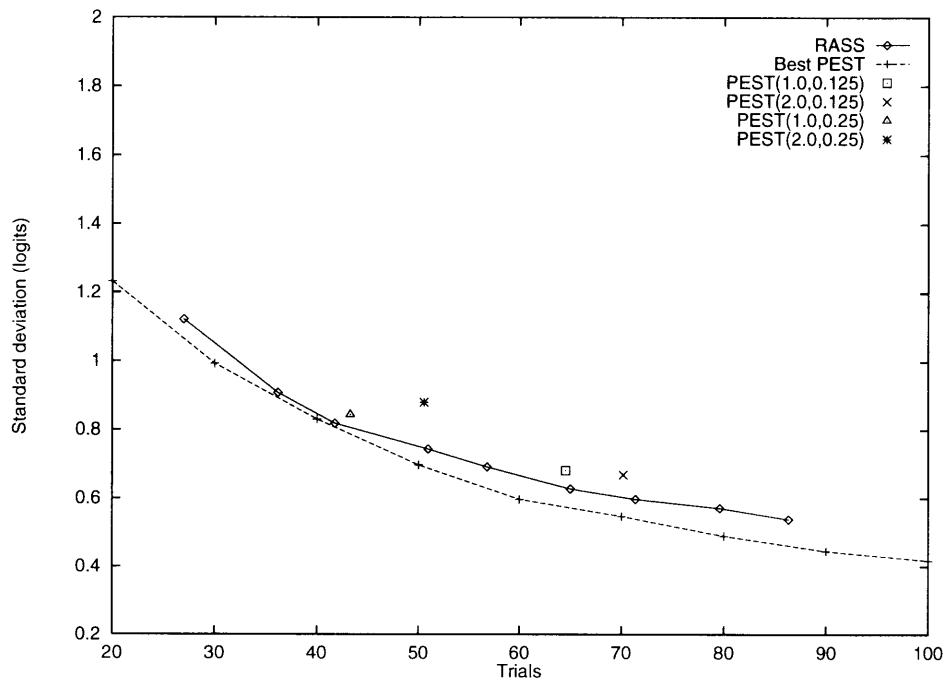


図 A.8: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 9(logits) である。)

表 A.7: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 9(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 27.0 | 36.2 | 41.7 | 51.0 | 56.8 | 65.0 | 71.4 | 79.6 | 86.3 |
| Standard deviation | 7.1 | 8.9 | 8.9 | 10.8 | 10.9 | 12.2 | 12.6 | 13.6 | 14.0 |

表 A.8: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 9(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 64.5 | 27.7 |
| PEST(2.0,0.125) | 70.2 | 33.6 |
| PEST(1.0,0.25) | 43.3 | 19.2 |
| PEST(2.0,0.25) | 50.5 | 23.6 |

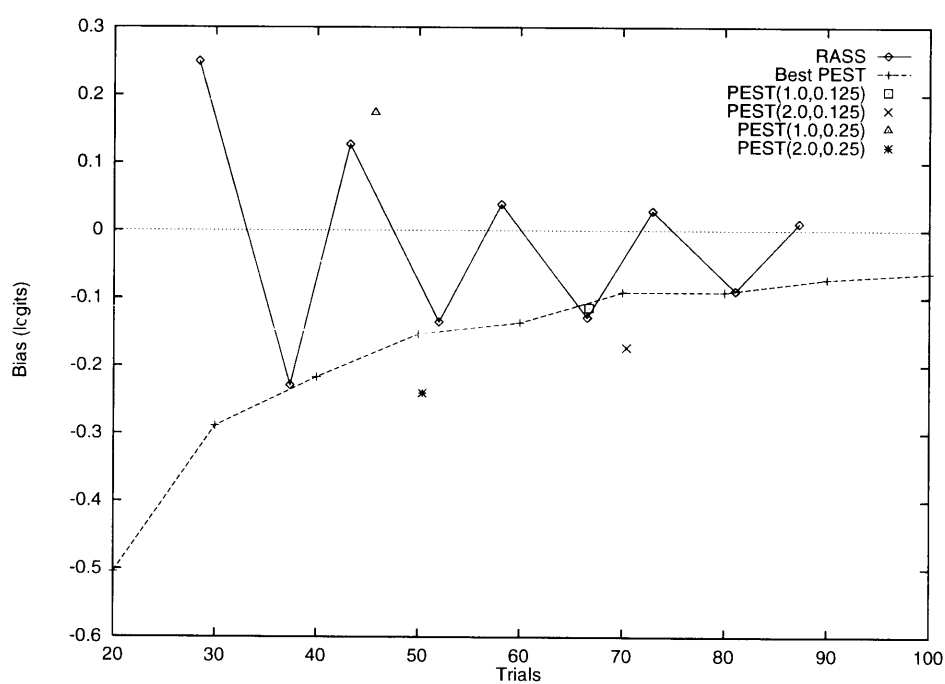


図 A.9: 3種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 10(logits) である。)

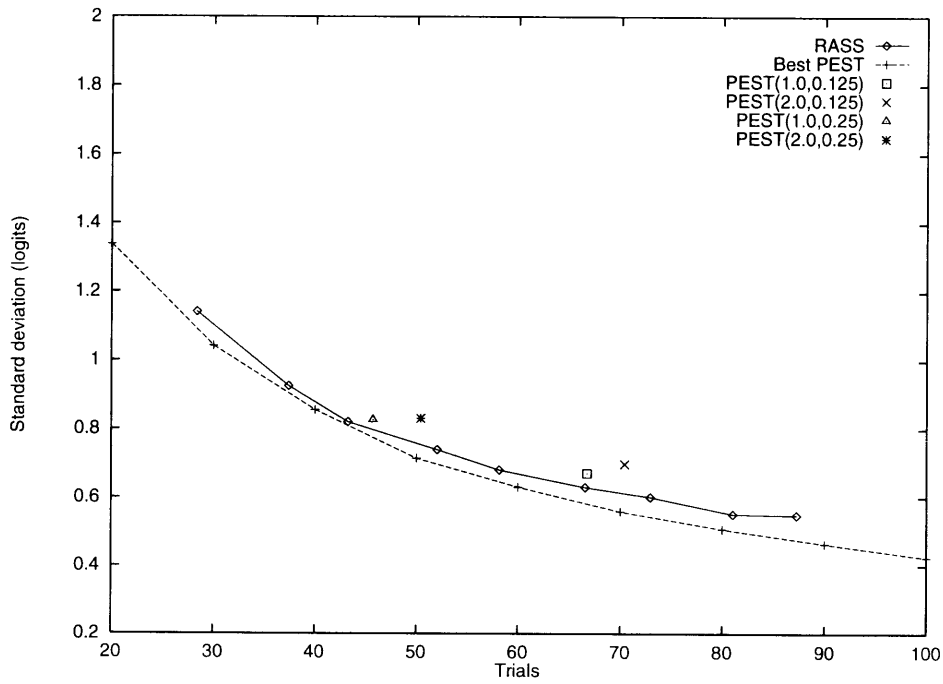


図 A.10: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 10(logits) である。)

表 A.9: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 10(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 28.4 | 37.4 | 43.2 | 52.0 | 58.1 | 66.6 | 72.9 | 81.1 | 78.3 |
| Standard deviation | 7.0 | 8.6 | 9.1 | 10.4 | 11.0 | 12.4 | 12.8 | 13.8 | 14.2 |

表 A.10: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 10(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 66.7 | 26.9 |
| PEST(2.0,0.125) | 70.4 | 33.2 |
| PEST(1.0,0.25) | 45.7 | 19.3 |
| PEST(2.0,0.25) | 50.4 | 23.5 |

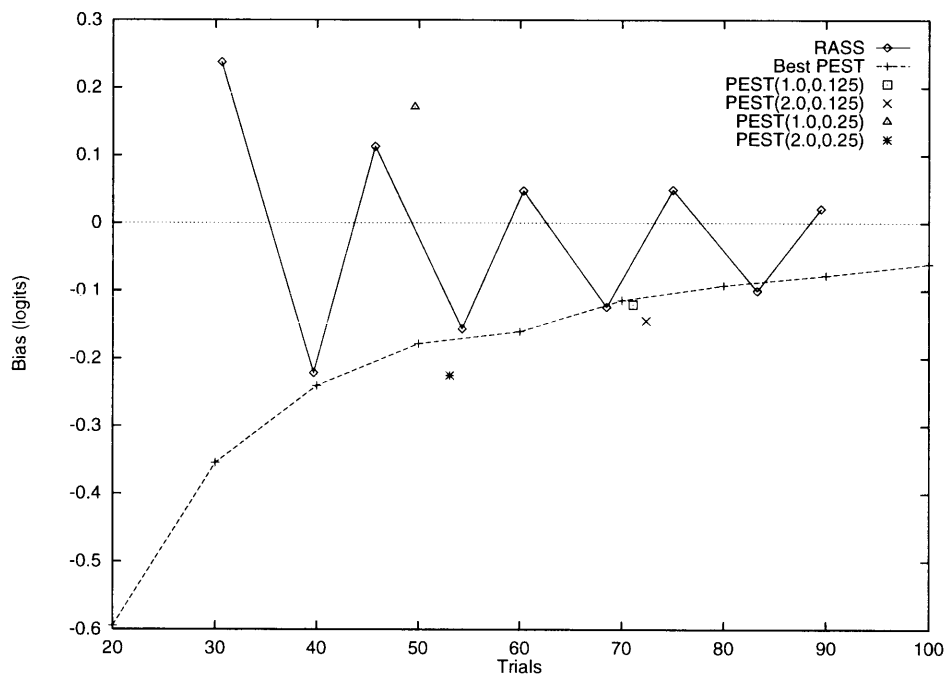


図 A.11: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 12(logits) である。)

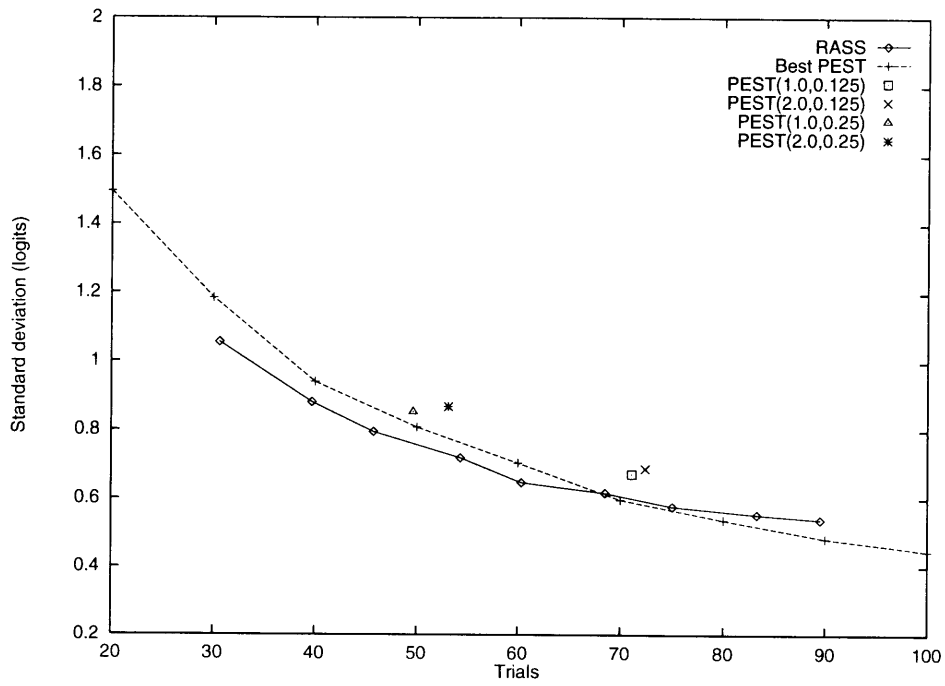


図 A.12: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 12(logits) である。)

表 A.11: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 12(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 30.6 | 39.7 | 45.7 | 54.3 | 60.3 | 68.5 | 75.0 | 83.3 | 89.5 |
| Standard deviation | 7.0 | 8.6 | 9.1 | 10.7 | 11.0 | 12.1 | 12.7 | 13.4 | 14.2 |

表 A.12: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 12(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 71.1 | 27.6 |
| PEST(2.0,0.125) | 72.4 | 35.5 |
| PEST(1.0,0.25) | 49.6 | 19.3 |
| PEST(2.0,0.25) | 53.1 | 23.6 |

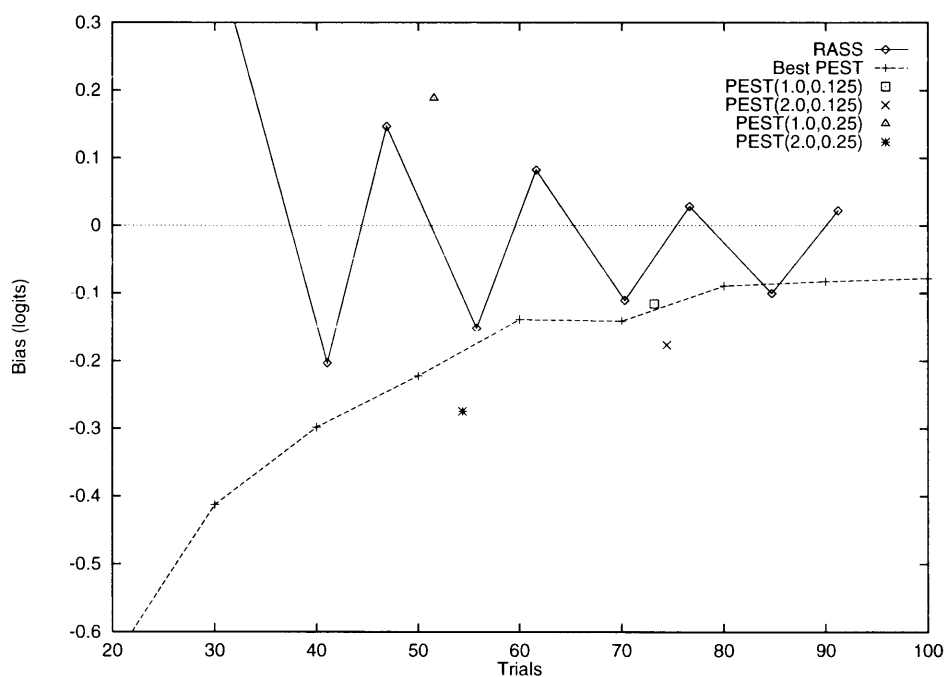


図 A.13: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 13(logits) である。)

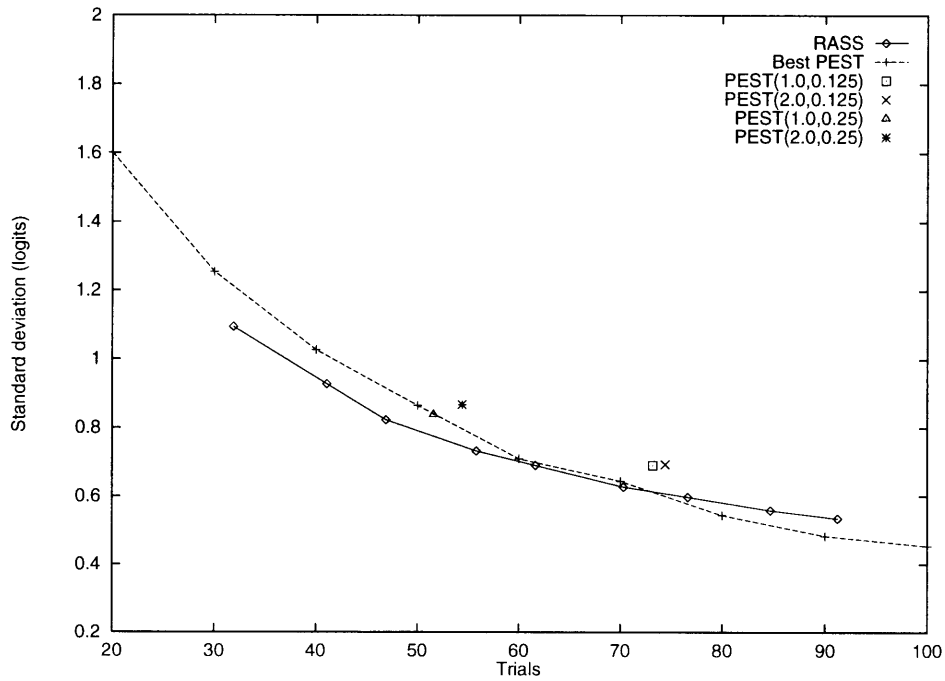


図 A.14: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 13(logits) である。)

表 A.13: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 13(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 31.9 | 41.1 | 46.9 | 55.8 | 61.6 | 70.3 | 76.6 | 84.7 | 91.2 |
| Standard deviation | 7.1 | 8.6 | 8.9 | 10.4 | 10.9 | 12.2 | 12.9 | 13.7 | 14.2 |

表 A.14: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 13(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 73.2 | 27.4 |
| PEST(2.0,0.125) | 74.4 | 34.9 |
| PEST(1.0,0.25) | 51.5 | 19.7 |
| PEST(2.0,0.25) | 54.4 | 23.7 |

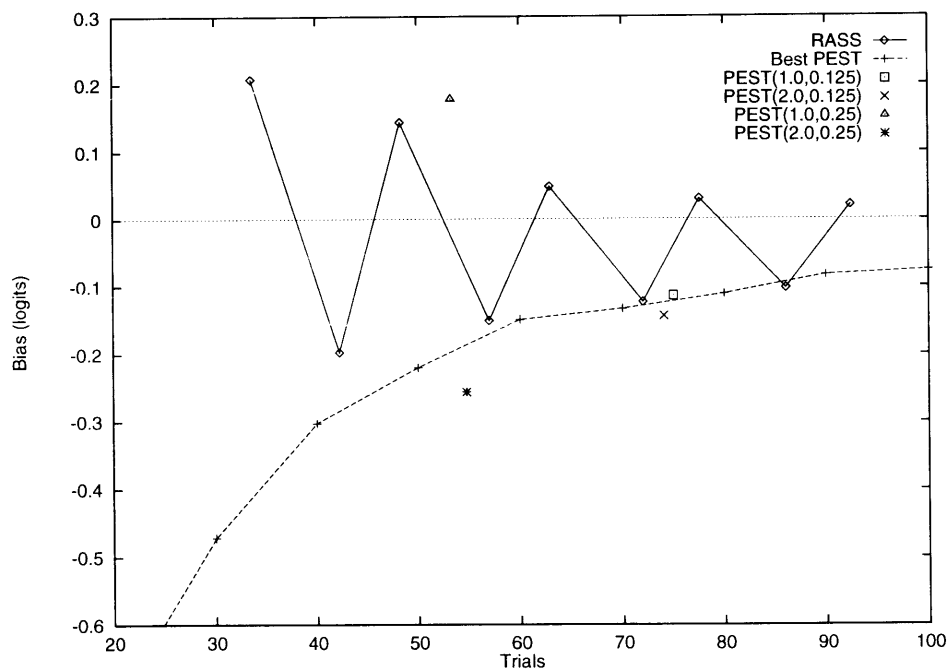


図 A.15: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 14(logits) である。)

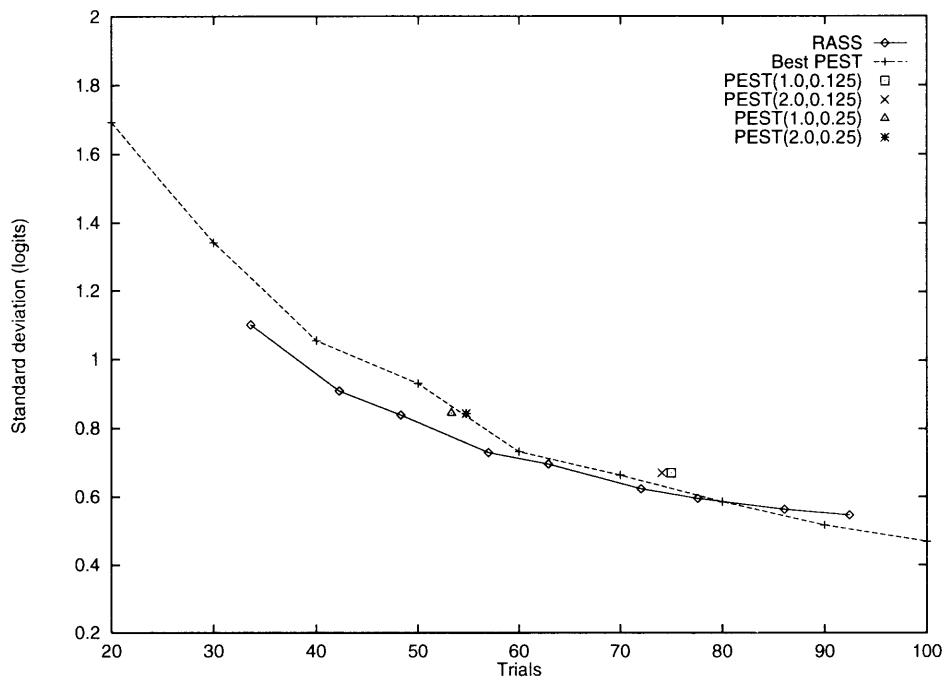


図 A.16: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 14(logits) である。)

表 A.15: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 14(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 33.6 | 42.3 | 48.3 | 57.0 | 62.9 | 72.0 | 77.6 | 86.1 | 92.4 |
| Standard deviation | 7.2 | 8.6 | 9.2 | 10.5 | 10.8 | 12.5 | 12.7 | 13.9 | 14.1 |

表 A.16: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 14(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 75.0 | 27.0 |
| PEST(2.0,0.125) | 74.1 | 33.6 |
| PEST(1.0,0.25) | 53.3 | 19.2 |
| PEST(2.0,0.25) | 54.8 | 23.3 |

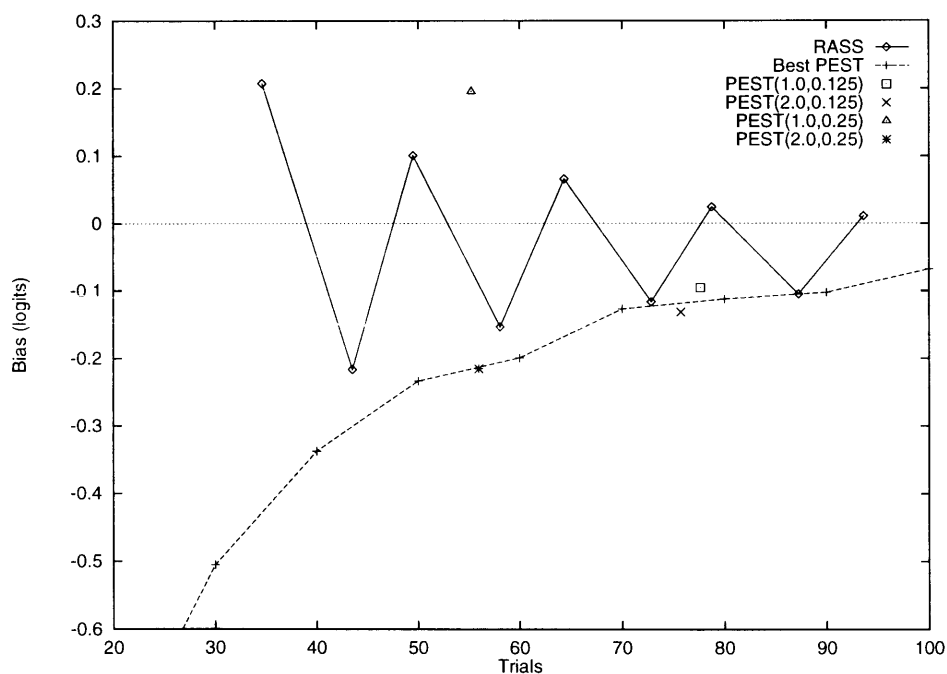


図 A.17: 3 種類の適応法の閾値推定のバイアス (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 15(logits) である。)

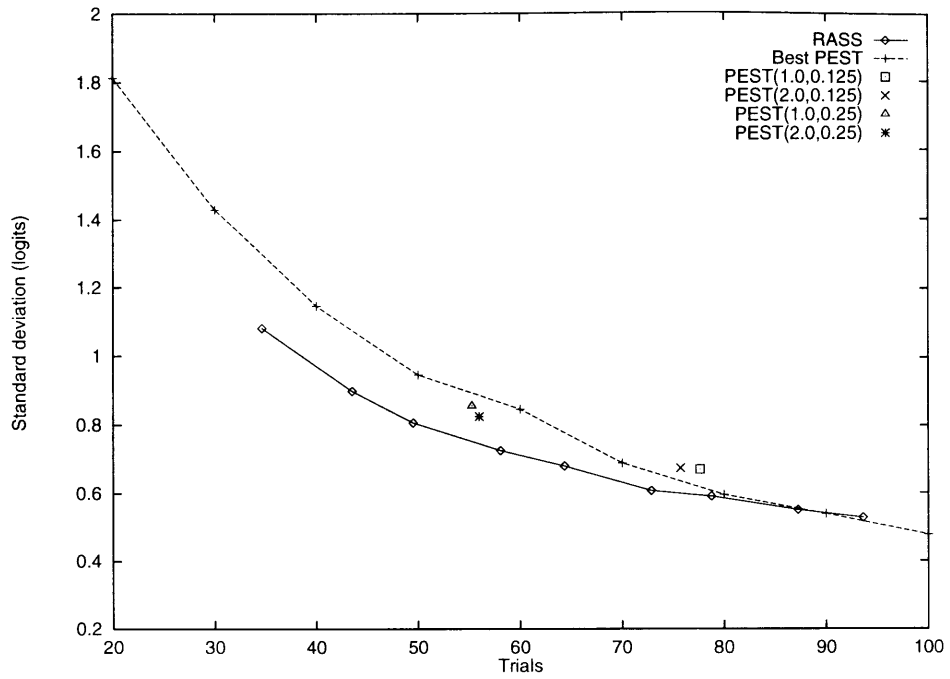


図 A.18: 3 種類の適応法の閾値推定の標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 15(logits) である。)

表 A.17: RASS 法の平均試行数及び標準偏差 (山・谷の個数は刺激レベルの変化方向が逆転した個数である。刺激レンジは 15(logits) である。)

| 山・谷の個数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------|------|------|------|------|------|------|------|------|------|
| Mean trial number | 34.6 | 43.5 | 49.5 | 58.1 | 64.4 | 72.9 | 78.8 | 87.3 | 93.6 |
| Standard deviation | 7.2 | 8.8 | 9.1 | 10.6 | 11.3 | 12.0 | 12.5 | 14.1 | 14.3 |

表 A.18: PEST 法の平均試行数及び標準偏差 (PEST(x,y) は x が最大ステップサイズ、y が stopping ステップサイズを示す。刺激レンジは 15(logits) である。)

| | Mean trial number | Standard deviation |
|-----------------|-------------------|--------------------|
| PEST(1.0,0.125) | 77.6 | 27.9 |
| PEST(2.0,0.125) | 75.7 | 35.2 |
| PEST(1.0,0.25) | 55.2 | 18.6 |
| PEST(2.0,0.25) | 55.9 | 23.7 |