

# 信号検出理論について — 数式と実践

石田 翼

2001年8月2日

## 1 初めに

Fechner 以来の精神物理的測定法の分野の一つとして、閾値の測定が挙げられる。閾値によって人間の感覚の感度を測ろうというのだ。

しかし Tanner & Swets (1954) は信号検出理論 theory of signal detection (detectability)/signal detection (detectability) theory をもちいることによって、閾値に拠らずに人間の感覚の感度を測定した。

信号検出理論は、1950年代にレーダー・システムの通信工学的理論として考案された(竹内ら編, 1989, p.804)。そもそもは信号 signal とノイズ noise との弁別力を調べ、レーダーの性能評価をしようとして作られたようだ。つまり、検出すべき信号はノイズの中にある、という状況で用いられるものであり、ノイズだけの状態と、ノイズの中に信号があるという二つの状況を弁別する能力を計測するのが目的である。精神物理学的測定法に応用するには、ノイズと信号を、無視すべき刺激と知覚すべき刺激とにそれぞれみなし、被験者の弁別力を算出するのだ。

この理論は実験心理学では基本的な手法の一つであり、それゆえ様々な本で理論は紹介されている。しかし、実際に使用する際の注意点や、測度が  $d'$  と  $A'$  の2種類あるがその違いは何か、などといった実用的な事について書かれたものは少ない。そこで、この文章でそれらの事について説明したいと思う。具体的には、実際に使用する際に必要な数式と理論を中心に説明していく。

まず最初に以降の説明に必要な信号検出理論の理論を説明し、それらの基本的な利用法を解説する。その後その利点と欠点を説明し、その欠点の改善を図った理論を説明する。最後に記憶の分野における信号検出理論の応用について述べる。

ただし本稿は実践を目的としているため、理論的に立ち入った話、例えば高閾値説 high threshold theory・低閾値説 low threshold theory・理想的受信者 ideal receiver の話など、には触れない。これらの話題について詳しく知りたいならば、Coombs, et al. (1970) や Gescheider (1985) などを参考のこと。

## 2 信号検出理論とは

### 2.1 $d'$ について

まず幾つかの仮定が必要である。物理的な刺激は被験者の内部で、その強度に対応した1次元の心理量に変換されるとする。そしてノイズだけの試行と、ノイズと刺激の両方を呈示される試行では、後者の方が物理エネルギー量が高い分、得られる心理量の平均値も大きいとする。また同等の刺激であっても、被験者の内部で常に同じ心理量を与えられるかというところではなく誤差があり、全体としてはその誤差は正規分布をとるとする。そしてノイズのみの正規分布も、ノイズ+刺激の正規分布も、同じ分散を持つと仮定する。

これらの仮定を図で説明するならば図1のようになる。図1横軸は1次元の心理量であり、ノイズ+刺激の正規分布の方が心理量が大きい分、右によっている。一方縦軸は頻度である。それぞれの正規分布の平均値が、物理量に対応した心理量であるので、その値を得る頻度がもっとも多く、平均値からずれるにしたがって頻度は減っていく。

さらに被験者は、刺激の強度がある一定以上の場合は刺激がある、と反応し、それ以下の場合は刺激がない、と反応するとする。その一定の基準 criterion が、図1に示す点線である。

そうすると、刺激+ノイズの正規分布のうち、点線よりも右側の部分の面積は、刺激+ノイズの試行で刺激ありと答えた正解（ヒット hit と呼ばれる）の確率と等しくなり、一方ノイズのみの正規分布のうち点線よりも右側の部分の面積が、ノイズのみの試行で刺激ありと答える間違った（フォールスアラーム false alarm と呼ばれる）確率である。ちなみに刺激+ノイズの試行で刺激なしと答える間違いはミス miss と呼び、ノイズのみの試行で刺激なしと答える正解はコレクトリジェクション correct rejection と呼ぶ。

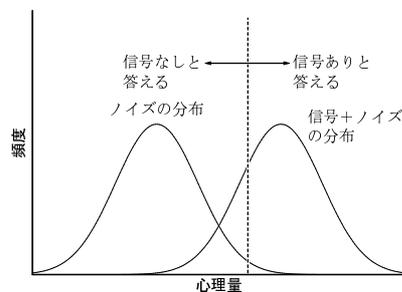


図 1: 信号検出理論

それら二つの確率は実験によって計測され、それらからこの基準からそれぞれの正規分布の平均値との間隔、ひいてはこの二つの正規分布の平均値間隔を算出することができる。これは物理量に対応した心理量の間隔であり、すなわち被験者のその刺激についての弁別力の指標である。これは  $d'$  で表される。それぞれの正規分布の分散を 1 とすると、この  $d'$  は分散が 2 の正規分布に従うことになる。基準からノイズ分布の平均までの距離を  $Z_N$ 、同じく刺激+ノイズ分布の平均までの距離を  $Z_{SN}$  とすると以下の式 (1) のように表される。

$$d' = Z_N - \frac{1}{S} Z_{SN} \quad (1)$$

### 2.1.1 $d'$ の分散

この  $d'$  は、式 (2) によって表される分散をもつ正規分布に従う (Gourevitch & Galanter, 1967, p.32)。

$$\hat{V}(d') = \frac{p_{FA}(1-p_{FA})}{n_N N(z_C)} + \frac{p_H(1-p_H)}{n_{SN} N(z_M)} \quad (2)$$

ここで  $p_{FA}$ : フォールスアラームの割合,  $p_H$ : ヒットの割合,  $n_N$ : ノイズ試行の試行数,  $n_{SN}$ : 刺激+ノイズ試行の試行数,  $z_C$ : コレクトリジェクションの割合の Z 得点,  $z_M$ : ミスの割合の Z 得点,  $N(z)$ : 正規分布である<sup>1</sup>。

### 2.1.2 バイアスの指標

さらにこの方法で被験者の反応の偏り、バイアス bias も調べられる。バイアスを調べるのは、その一定の基準値がどこにあるかであらず。もし被験者の反応が、ノイズのみの分布とノイズ+刺激の分布のどちらにも偏っていないならば、基準は二つの正規分布のちょうど真ん中、それらの交点にくる。しかしどちらかを多く答えるように偏っていると、この基準値はそちらの分布の方に偏った位置にあること

<sup>1</sup>この式から算出された分散は標本分散であるので、検定の際に用いる分散を(統計の教科書に良く出ている式にあるように)標本数で割るという必要はない。なぜならば、一般的に統計の入門書に出てくるような基本的なケースでは、標本の分散を母集団の不偏分散と見なしている。そして標本は正規分布を満たしている(中心極限定理により)。そのため、統計量を計算する際には標本数で割って、標本の正規分布の分散を算出しなければならないのだ。今回の場合この式から算出される分散はすでに標本分散であるので、標本数で割る必要はない。

になる．その偏り具合は，その基準値におけるそれぞれの正規分布の値の比によってあらわされる．先程の図 1 に則して説明するなら，それぞれの正規分布と基準値の点線との交点の，それぞれの高さの比である．これは  $\beta$  と呼ばれ，1 だとバイアスなしで，それより大きいと刺激があると答えがちなバイアスがあり，1 以下だとノイズのみと答えがちなバイアスがある，と解釈される．この  $\beta$  は自然対数を取ると，分散が  $d'^2$  の正規分布に従う．

### 2.1.3 利点と欠点

先程も述べたとおり，人間の感覚の感度を調べることによって，そのメカニズムを探ろうというのが実験心理学の重要な手法の一つである．この感度の測度には，閾値が使われることが多かったが，この信号検出理論は閾値とはまた別の考え方による感度の測定法である．

閾値の測定という方法と比べると，この方法は被験者のバイアスと，それを取り除いた弁別力とを同時に計測できる点で優れている．また，S. S. Stevens によると (cited in Gescheider, 1985, p.132)，ノイズが多い状況だと閾値というものが疑わしくなる<sup>2</sup>ので，信号検出理論の方が好ましいようだ．

また，この理論は，今まで精神物理学分野では閾値によって成績を測定されていたような，広く一般のノイズ・刺激の弁別課題に応用できる．またさらにこのノイズと刺激という語の解釈によっては様々な分野に応用できる．例えば，記憶研究での再認の成績を測るのにも使われている (Pollack & Norman, 1964; Pollack et al., 1964; Banks, 1970; Lockhart & Murdock, 1970 など)．このことに関しては，後に章を改めて述べる．また他には痛みの研究や臨床場面でも使われているという (Gescheider, 1985, pp.123-132)．

次に欠点を見てみよう．

まず  $\beta$  についてだが，これは  $d'$  の値によって取り得る値が変わってくるのだ．例えば  $d'=0.0$  の時は  $\beta$  は常に 1 であるが， $d' = 1.0$  ならば 0.0 から約 10.0 までの値を取る．そして  $d' = 3.0$  となると， $\beta$  は最大約 100 の値を取り得るのだ (Banks, 1970, p.86)．これは先程述べた， $\beta$  の母集団の分散が  $d'$  の自乗であることにも現れている．このため，被験者の感度が異なれば，それぞれのバイアスを表す  $\beta$  の値が同じでも比較はできない．

これを改善するため，Banks (1970, p.86) はその代替となる測度  $C_j$  を提案している．これはノイズ分布の平均から，被験者の反応の基準までの距離である．これなら被験者の感度の影響もないし，単位が  $d'$  と同じ標準偏差で，先程の (2) 式をそのまま利用できる．しかし， $\beta$  にせよ  $C_j$  にせよ，この単位が標準偏差であるという点は欠点でもある．これはノイズ分布並びに刺激 + ノイズ分布の標準偏差であるが，これが常に一定か？という問題があるのだ．つまり，この標準偏差は被験者間や刺激間で変動するおそれがあるのだ．そうすると  $C_j$  を用いたとしても， $\beta$  の場合よりも問題は減るにせよやはり比較が可能か疑問が残る．

また，上記の議論は，あくまで正規性の仮定と等分散の仮定が満たされたと仮定しての話である．もし等分散性が満たされていないならば， $\beta$  が 1 になる点が複数出てくるので，このバイアスの指標が信用できなくなる (大山ら編，1994，p.41)．

おそらくその正規性の仮定を保証するためだと思われるが，Gescheider (1985, p.113) は信号検出理論を用いる際には試行は 300 回以上が望ましいと述べている．田中・上村 (1969, p.134) によれば 500 試行以上が典型的である．それより少ない試行数なら，これらの仮定が満たされていることを，あらかじめ確かめておく必要があるだろう．

以下でその確認の方法と，それらの仮定の性質，そしてそれらの仮定をせずに，つまりノンパラメトリックに同様な測定値を算出する方法を述べる．

<sup>2</sup>閾値という概念では，刺激にもノイズが含まれているということは考慮されていない (竹内ら編，1989, p.806)．

### 3 正規性並びに等分散性の仮定の確認方法

#### 3.1 ROC 曲線

これについては、ROC (receiver operating characteristic) 曲線 (受信者動作特性曲線・等感受性曲線とも言う) を調べるという方法がある。ある  $d'$  に対して、フォールスアラームの割合を  $x$  軸にとり、ヒットの割合を  $y$  軸にとると、 $d'$  に応じて曲線が描かれる。これが ROC 曲線である。その例を図 2 に示す。それぞれの  $d'$  について、ROC 曲線が描かれている。これら線上の、被験者の反応にかかっているバイアスに対応した位置に成績がプロットされる。したがって得られたデータが先の仮定を満たしているならば、データをプロットすればその曲線上に乗るのだ。

この ROC 曲線を描くには大きく 2 種類の方法がある。

1 つは何種類かのバイアスがかかる状況 (刺激 + ノイズの試行での正解の報酬とノイズのみの試行の正解での報酬とに差をつけたり、それぞれの出現確率を変えたりする) でデータを取り、それをプロットして ROC 曲線に乗るかどうかみてみればよい。

もう 1 つの方法は、被験者の反応を評価によってとる方法である (Pollack & Decker, 1958)。被験者の反応カテゴリーを yes と no の 2 段階にはせず、例えば「確実に (刺激) ある」「あるようだ」「全く分からない」「ないようだ」「全くない」の 5 段階の評価をさせるのだ。「確実にある」を 5 とし、「全くない」を 1 とすると、この評価値は被験者内の心理量を反映したものだと思えることができる。これをさらに上で説明したような一般的な方法のように、2 段階の yes/no に分割し直すとしてしよう。この時の分割の基準は、先に説明した被験者の反応を決定する基準値と同じものと見なせる。つまり、評価を 2 分する基準を操作することは、被験者のバイアスを擬似的に操作することになるのだ。

具体例を示そう。例えば表 1 のような結果を得たとしよう。横軸は被験者の評価値で、縦軸は刺激を提示したかしないかであり、表内の数値はそれぞれの評価値において、被験者が yes と反応した割合である。まず縦軸のカテゴリそれぞれでそれらの累計確率を求める。これが表である。そうするとこれは、yes/no に分割し直す基準をそれぞれの評価値の間に置いた場合のヒットとフォールスアラームの値となる。つまり、被験者の基準値が 5 と 4 の間、4 と 3 の間、3 と 2 の間、2 と 1 の間であるように操作した場合のそれぞれの成績である。これらからそれぞれの成績を再構成すると表ようになる。これらがそれぞれ ROC 曲線上の 4 つの点となるので、ROC 曲線を書けるのだ。一般に評価が  $n$  段階であれば、 $n - 1$  個の点が打てることになる。

この方法で書かれた ROC 曲線を、タイプ 2 の ROC 曲線 type II ROC curve と呼ぶ (Lockhart & Murdock, 1970)。一方上述の一般的な方法で求めた ROC 曲線はそれに対しタイプ 1 の ROC 曲線 type I ROC curve と呼ぶ。

ただしこの方法を用いるためには全てのセルの値が 0 以上でなければならないので、被験者の反応が偏っていない必要がある。

どちらの方法でも ROC 曲線を書いたにしても、実際は曲線では判断しにくいので、ヒットの割合とフォールスアラームの割合とを、 $Z$  得点化しプロットする。そうすると、ROC 曲線は直線になるので、直線回帰が使える。

その時に等分散性が満たされていれば回帰直線の傾きが 1 を示す。正規性が満たされているかどうか

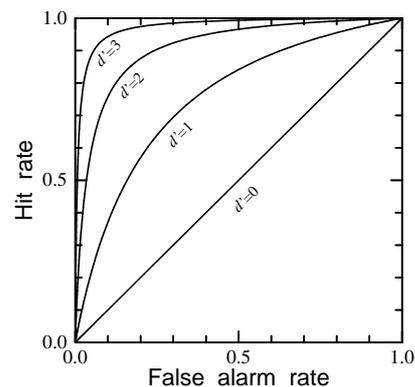


図 2: ROC 曲線の例

表 1: 評定から ROC 曲線を求める (例)

	刺激	刺激 + ノイズ
5	0.26	0.05
4	0.24	0.10
3	0.20	0.15
2	0.15	0.20
1	0.15	0.50

(a) 評定値データ

	刺激	刺激 + ノイズ
5	0.26	0.05
4	0.50	0.15
3	0.70	0.30
2	0.85	0.50
1	1.00	1.00

(b) 評定値データの累計

	刺激	刺激 + ノイズ
刺激あり反応 と見なす	0.26	0.05
刺激なし反応 と見なす	0.74	0.95

(c) 基準値を評定値 5 と 4 の間として得られた成績

	刺激	刺激 + ノイズ
刺激あり反応 と見なす	0.50	0.15
刺激なし反応 と見なす	0.50	0.85

(d) 基準値を評定値 4 と 3 の間として得られた成績

	刺激	刺激 + ノイズ
刺激あり反応 と見なす	0.70	0.30
刺激なし反応 と見なす	0.30	0.70

(e) 基準値を評定値 3 と 2 の間として得られた成績

	刺激	刺激 + ノイズ
刺激あり反応 と見なす	0.85	0.50
刺激なし反応 と見なす	0.15	0.50

(f) 基準値を評定値 2 と 1 の間として得られた成績

は、データプロットの直線へのあてはまり具合、すなわち相関係数から判断する。それによって先述の2つの仮定を検証できる。以下でこの2つの仮定が満たされていない場合の対応策を述べる。

### 3.2 等分散性が満たされていない場合

上記の検証の結果、等分散性の仮定が満たされていない<sup>3</sup>と分かったとしても、得られたROC曲線からそれを補正する測度が幾つか考えられているので、それを用いることが出来る。

#### 3.2.1 $d'$ を補正する方法

1つは正規化したROC曲線の傾きが、刺激+ノイズ分布とノイズ分布の分散の比と等しいことを利用する(Gescheider, 1985, pp.100-106)方法である。つまり実験内で差の少ないと思われるノイズ分布<sup>4</sup>の分散を基準として補正された $d'$ を算出するのだ。先程同様基準からノイズ分布の平均までの距離を $Z_N$ 、刺激+ノイズ分布の平均までの距離を $Z_{SN}$ とし、傾きを $S$ として式で表すと、以下の式(3)のようになる。

$$d' = Z_N - \frac{1}{S}Z_{SN} \quad (3)$$

#### 3.2.2 正規分布以外の分布を使う方法

正規化したROC曲線の傾きが1ではなく、しかも $d'$ の値にしたがって変化しているようであれば、刺激+ノイズとノイズの分布形を正規分布以外の分布にすると、当てはまりが良くなるということがある(Lockhart & Murdock, 1970, p.106)。例えばポアソン分布・対数分布・ガンマ分布等である。これらの分布は分散が平均値の関数になっているので、正規化したROC曲線の傾きが $d'$ の値にしたがって変化するのを補正することになるからだ。

#### 3.2.3 $\Delta_m$ を使う方法

$\Delta_m$ という指標を用いる方法もある(御領, 1973, p.67; 田中・上村, 1969, p.122)。この指標は、ヒットが0.5であるとき(Z得点で表すならば0であるとき)のフォールスアラームの値のZ得点であり、大抵は傾きを併用して記述する。先程同様傾きを $S$ で表すと $D(\Delta_m, S)$ と記述する。具体的には $D(\Delta_m, S) = 1.2, 0.8$ というように表す。

#### 3.2.4 ROC曲線から算出される指標 $d'_e$

また、ROC曲線と負の対角線との交点のZ得点を基に算出された $d'$ を指標に使うこともある。これの利点はROC曲線から、指標を算出できることにある。 $d'_e$ で表すとすると、先述の $\Delta_m$ とは式(4)で表せる関係を持つ。

$$d'_e = 2\Delta_m \left( \frac{S}{1+S} \right) \quad (4)$$

この指標の性質は $d'$ に準ずる。

#### 3.2.5 強制選択法の使用

<sup>3</sup>この等分散性が満たされないという状況は、被験者が刺激について十分な知識を持たず、情報を十分に用い得ない場合に見られるようだ。この様な状況は視覚実験において多く見られる(Swets, 1961 cited in Coombs, et al., 1970, 翻訳書 p.190)。逆に被験者が信号に含まれている情報を十分に利用できる時には等分散性が見られるようだ(田中・上村, 1969)。

<sup>4</sup>実験で操作されるのは一般に刺激の方であるので、操作されないノイズの方は実験内では変動しないと見て良い。

他には、測定時の手法として強制選択法を用いると、この方法の仮定として等分散性が満たされることになる（御領，1973，p.69）。強制選択法とは、複数の試行を被験者に提示するのだが、その中のいずれか1つだけに刺激が含まれていて、被験者には刺激が含まれていたと思われる試行を1つ答えてもらうというものだ。複数の刺激は時間的または空間的位置をずらして提示される。

この方法の利点は、被験者は常にいずれかの提示に刺激が含まれていると知っているので、反応の基準が実験内であまり変化しないという点と、前述のように、等分散性が得られる点である。また、一般的な yes/no という選択肢と比べると、提示されたどれを選ぶか、というこの方法における選択の方が当然バイアスはかかりにくい、という利点もある（Gescheider, 1985, pp.117）。

### 3.3 正規性の仮定

先に、正規性が満たされているならばZ得点化したROC曲線が直線になる、と述べたが、実はこの逆は成り立たないことが観察されている。つまり、Z得点化したROC曲線が直線になったからといって、正規分布であるとは限らないのだ。Lockhart & Murdock (1970, pp.105-106) がノイズの分布を矩形の分布に、刺激 + ノイズの分布を三角形の分布にしてROC曲線を書いたところ、ほぼ直線になった。

これを再現した図を図3に示す。これらの、明らかに正規分布とはかけ離れた分布からROC曲線を書いても、図3のように直線になることが分かる。

確かに正規分布の前提からは、ROC曲線をZ得点化すると直線になる、という予測は得られる。しかし正規分布以外の分布でも直線になり得るのだ。これはROC曲線は分布の形にあまり影響を受けないという頑健性でもあるが、同時にROC曲線から分布の形を検証することが出来ない、という欠点でもある。

また中心極限定理のため、想定される分布がどんな形であれ標本分布は正規分布になってしまう。これらのことから、正規性の仮定についてはあまり気にする必要はない。

## 4 ノンパラメトリックな信号検出理論

### 4.1 $A'$ の算出

上記のような検証ができず、正規性・等分散性を仮定できないような場面では、それらの仮定がなくともできるようなノンパラメトリックな手法が考案されている。それを以下で説明する。

図4において、横軸はフォールスアラームの、縦軸はヒットの、それぞれ割合である。図中に示された  $x, y$  は測定されたそれぞれの値である。するとここで  $(1, 1)$  の点からと  $(0, 0)$  の点から  $(x, y)$  を通る直線が引ける。一方ROC曲線は  $(x, y)$  の点を含むので、 $A_1$  と  $A_2$  の領域内を通る。よって  $I$  は明らかにROC曲線よりも下の領域であり、 $A_1$  と  $A_2$  はROC曲線よりも上の領域になるか下の領域になるか明らかではない領域である。

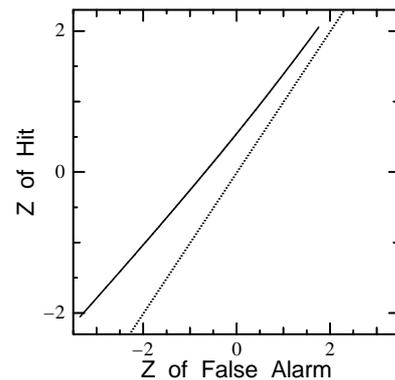


図 3: 正規分布以外の分布による ROC 曲線の例。実線で表されているグラフが本文にもある分布関数から算出された ROC 曲線。点線で表されている直線は、比較のためにかかれた、 $d' = 0$  の ROC 曲線。

ところで、このような単位正方形の中での ROC 曲線よりも下側の領域の面積は、バイアスを取り除いた、修正された正解率に等しい (Green, 1964; 田中・上村, 1969, p.112-116)。そこで、 $A_1$  と  $A_2$  の面積の半分と  $I$  の面積との和を ROC 曲線より下の面積と仮定すれば、それはこの場合のバイアスのかかっていない正解率と見られる (Norman & Galanter, 1964; Polack & Norman, 1964)。これは  $A'$  と呼ばれているもので、数式で表すと以下の式 (5) のようになる。

$$A' = I + \frac{A_1 + A_2}{2} \quad (5)$$

一般には  $x$  と  $y$  で書き直した (6) のような形で使われている (Grier, 1971)。

$$A' = \frac{1}{2} + \frac{(y-x)(1+y-x)}{4y(1-x)} \quad (6)$$

#### 4.1.1 $A'$ におけるバイアスの算出の仕方

ではこの場合のバイアスの算出法はどうであろうか。Hodos (1970) の方法を Grier (1971) が改良したものを紹介しよう。この方法は先の図 4 で言うと、 $A_1$  と  $A_2$  の面積の比からバイアスを求める方法である。具体的に数式にすると式 (7) のようになる。

$$B'' = \frac{A_1 - A_2}{A_1 + A_2} \quad (7)$$

$x$  と  $y$  で書き直すと以下の式 (8) のようになる。

$$B'' = \frac{y(1-y) - x(1-x)}{y(1-y) + x(1-x)} \quad (8)$$

#### 4.2 ノンパラメトリックな手法とオリジナルな手法の違い

これらノンパラメトリックな測度は、オリジナルの測度である  $d'$  や  $\beta$  とは異なった性質を持つ。 $d'$  は仮定から正規分布にしたがうと見て良い。一方  $A'$  は 0 以上 1 以下の値をとる補正された正解率であり、分布の形は二項分布に従うと見なして良い (Pollack & Hsieh, 1969, p.166)。また、 $\beta$  は 1 の時にバイアスがなく、最大値・最小値はそれぞれ無限大・無限小であるのに対し、 $B''$  は 0 の時バイアスなしで最大値・最小値はそれぞれ  $+1$ ・ $-1$  である。

## 5 まとめ

### 5.1 どの測定方法を使うべきか

Macmillan & Creelman (1991, cited in Donaldson, 1993) は、「評価によらないで信号検出理論を用いるのは危険だ」と述べている。なぜならば、普通に yes/no の二件評価で算出した値は、ROC 曲線のある一点のみを算出することであり、そこからある測度を導き出すのは、その一点だけから ROC 曲線全体を導き出すのに等しいからだ。つまり  $d'$  は ROC 曲線に対応しているし、 $A'$  は ROC 曲線の下側の範囲の面積であるので。

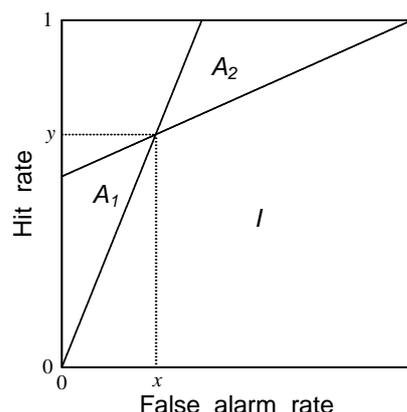


図 4: ノンパラメトリックな信号検出理論

これは、一般化としてはいささか過剰である。したがって評定によらない信号検出理論の測度は信頼性にかける。可能な限り、被験者の成績の評価には、評定によって ROC 曲線を描いたうえで測度を算出するという方法を取るべきだろう。ROC 曲線さえ測定しておけば、仮定が成り立っているかどうかも分かるし、成り立っていないとしてもそれを補正できる。

それが出来ない場合、次善の策としては強制選択法を用いることである。

しかしこれら 2 つの方法はバイアスを測ることが出来ない。ROC 曲線はバイアスとは無関係に描かれるものであるし、強制選択法では理論上バイアスが発生しにくいからだ。バイアスを測りたいのであれば、オリジナルな yes/no の 2 択による指標を用いるしかない。

## 5.2 どの測度を使うべきか

そのような yes/no の 2 択、あるいは強制選択法を用いた場合は、 $d'$  と  $A'$  のどちらの測度を用いるべきであろうか。

測度として  $d'$  を使うオリジナルの信号検出理論は、参考書にもよく説明があり、一般的である。しかし先述したようにこれは前提となる条件が多く、その条件が保証されている事を確認するのも手がかかる。

さらに Donaldson (1993) が、分布の平均間の距離やバイアス様々に変化させて  $d'$  と  $A'$  の精度を調査したところ、一般的に  $d'$  よりも  $A'$  の方が正確である、という知見を得た。ただし等分散性が満たされ、バイアスもかかっていないという場合には  $d'$  の方が正確であるようだ。

したがって ROC 曲線を算出した上で指標を算出するのであれば  $d'$ 、それ以外の場合は  $A'$  を使用するのが適切である。

実際に  $d'$  を算出する場合には、積分した正規分布の逆関数が必要であり、これを計算するのは代数的には困難である。しかし幾つか近似式が考えられており、それで代用できる（これについては、付録で紹介している）。

## 参考文献

- [1] Banks, W. P., 1970, Signal detection theory and human memory. *Psychological Bulletin*, 74, 81-99.
- [2] Coombs, C. H., Dawes, R. M., and Tversky, A., 1970, *Mathematical psychology: An elementary introduction*. Prentice-Hall. (小野 茂監訳, 1974, 数理心理学序説. 新曜社)
- [3] Donaldson, W., 1993, Accuracy of  $d'$  and  $A'$  as estimates of sensitivity. *Bulletin of the Psychonomic Society*, 31, 271-274.
- [4] 御領 謙, 1973, 認知閾の測定と信号検出理論の応用. (大山 正編, 1973, 心理学研究法第2巻 実験1. 東京大学出版会. 58-69 収録)
- [5] Gescheider, G. A., 1985, *Psychophysics: Method, Theory, and Application*. Lawrence Erlbaum Associates.
- [6] Gourevitch, V., and Galanter, E., 1967, A significant test for one parameter isosensitivity functions. *Psychometrika*, 32, 25-33.
- [7] Green, D. M., 1964, General prediction relating Yes-No and forced-choice results. *Journal of Acoustical Society of America*, 36, 1042(Abstract).
- [8] Grier, J. B., 1971, Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.
- [9] Hodos, W., 1970, Nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, 74, 351-354.
- [10] 情報システムハンドブック編集委員会編, 1989, 情報システムハンドブック. 培風館.
- [11] Lockhart, R. S., and Murdock, B. B. Jr., 1970, Memory and the theory of signal detection. *Psychological Bulletin*, 74, 100-109.
- [12] Murdock, B. B., Jr., 1968, Serial order effect in short-term memory. *Journal of Experimental Psychology monograph Supplement*, 76, part 2.
- [13] 大山 正・今井 省吾・和気 典二編, 1994, 新編感覚知覚ハンドブック. 誠信書房.
- [14] Pollack, I. and Decker, L. R., 1958, Confidence ratings, message reception, and the receiver operating characteristic. *The journal of the acoustical society of America*, 30, 286-292.
- [15] Pollack, I., and Hsieh, R., 1969, Sampling variability of the area under the ROC-curve and of  $d'$ . *Psychological Bulletin*, 71, 161-173.
- [16] Pollack, I., and Norman, D. A., 1964, A non-parametric analysis of recognition experiments. *Psychonomic Science*, 1, 125-126.
- [17] Pollack, I., Norman, D. A., and Galanter, E., 1964, An efficient non-parametric analysis of recognition memory. *Psychonomic Science*, 1, 327-328.
- [18] 篠崎 信雄, 1994, 統計解析入門. サイエンス社 (MSライブラリ - 1) .

- [19] 竹内 啓ら編, 1989, 統計学辞典. 東洋経済新報社.
- [20] 田中 良久・上村 保子, 1969, 信号検出理論. (田中 良久編, 1969, 講座心理学 第2巻 計量心理学. 東京大学出版会. 105-140 収録)
- [21] Tanner, W. P., Jr., and Swets, J. A., 1954, A decision-making theory of visual detection. *Psychological Review*, 61, 401-409.

## 付録

### A $d'$ を計算する際に必要な数式

本文において  $d'$  の算出の理論を解説した際に述べた通り、算出の際には正規分布のいわゆる上側確率から、 $Z$  得点、つまりパーセント点を算出する必要がある。これは正規分布式を積分した後に逆関数を求めればよいのだが、これは代数的に計算式を導出することは出来ない。入門書（例えば篠崎（1994））や統計数値表には算出された値が表の形で載っている場合があるが、実験場面などでその場ですぐに計算する必要がある場合には実用的ではない。

この目的のために幾つかの近似式が考案されているので、そのうちの一つ、戸田の近似式と呼ばれるもの（cited in 情報システムハンドブック編集委員会編, 1989, 3-210）をここで紹介する。上側確率を  $Q$  とし、パーセント点を  $u(Q)$  とすると、正規分布の上側確率が  $Q$  となる点  $u(Q)$  は以下の近似式で表される。

$$u(Q) = \left[ y (b_0 + b_1 y + b_2 y^2 + \dots + b_{10} y^{10}) \right]^{\frac{1}{2}} \quad (0 < Q \leq 0.5)$$

ただし

$$y = -\ln [4Q(1-Q)]$$

$$b_0 = 0.1570796288 \times 10$$

$$b_1 = 0.3706987906 \times 10^{-1}$$

$$b_2 = -0.8364353589 \times 10^{-3}$$

$$b_3 = -0.2250947176 \times 10^{-3}$$

$$b_4 = 0.6841218299 \times 10^{-5}$$

$$b_5 = 0.5824238515 \times 10^{-5}$$

$$b_6 = -0.1045274970 \times 10^{-5}$$

$$b_7 = 0.8360937017 \times 10^{-7}$$

$$b_8 = -0.3231081277 \times 10^{-8}$$

$$b_9 = 0.3657763036 \times 10^{-10}$$

$$b_{10} = 0.6936233982 \times 10^{-12}$$

相対誤差は  $1.2 \times 10^{-8}$  であるので、パーソナルコンピュータで倍精度実数で計算し単精度実数で出力する限り、丸め誤差に収まる。

### B Murdock (1968) の測定方法

Murdock (1968) が記憶における系列位置効果を見た実験において、信号検出理論を興味深いやり方で使っているので、その方法を Banks (1970) を参考に紹介する。彼の信号検出理論の使い方で注目すべきは、系列内の位置の記憶・再生の成績を、この理論で評価したことである。その方法は以下のとおりである。

記憶すべき系列中のある項目が提示された順番を  $i$  , それが再生された順番を  $j$  で表す . ここで  $i$  を行に ,  $j$  を列にして , 全ての項目について被験者の解答の頻度の表を作る . つまり  $i$  行  $j$  列のセルに ,  $i$  番目に提示された項目が  $j$  番目に再生された頻度を入れる .

ここで , 信号検出理論を適用するために , ある項目が提示されたときの系列位置と同じ再生時の系列位置を刺激と見なし , それ以外の再生時の系列位置をノイズと見なす . そこで  $i, i$  のセル内の頻度は , ヒット (表中 H) の頻度と見なせる . さらに  $i$  番目には提示された項目を  $i$  番目に再生し損ねるのは , 再生すべき系列位置を見逃した , すなわちミス (表中 M) と見なせる . また  $i$  番目に提示されていない項目を  $i$  番目に再生するのはフォールスアラーム (表中 FA) であり , それ以外のものはコレクトリジェクション (表中 CR) となる .

これらそれぞれの頻度を全体の総数で割って割合を求めれば , これはそのまま信号検出理論を適用できる . しかし残念ながらこの方法が適切かどうかの検証はなされていない . この方法の利点は , 再生の成績を信号検出理論で評価する際に , 被験者の評定以外の方法で測れるということである . 先述のように評定による信号検出理論はバイアスを測れないという欠点があるので , それを克服できる .

また , この方法は系列再生の成績を信号検出理論で評価することが可能である . この利点は , 記憶研究以外の分野でも応用の可能性があると思われる .

表 2: Murdock の方法  
再生時系列位置

	...	$j$	...	
提示時	⋮	CR	FA	CR
系列位	$i$	M	H	M
置	⋮	CR	FA	CR