

原 著

多変量解析による虚偽検出検査の自動判定

足 立 浩 平

AN AUTOMATED DIAGNOSTIC METHOD FOR DETECTION OF DECEPTION USING MULTIVARIATE STATISTICAL ALGORITHM

Kohei ADACHI

A computer-based diagnostic system is developed for the guilty knowledge test. The system measures and analyzes magnitudes of physiological responses to questions, to judge whether a subject has guilty knowledge or not. Measures of response magnitudes are respiration speed, SCR amplitude, baseline change of pulse wave, and heart rate change defined using the auto-regression model. These four variables are analyzed by multivariate statistical algorithm which we propose.

The statistical algorithm is basically to compare validity between an innocent and a guilty hypotheses. The innocent hypothesis assumes that a subject's response magnitudes for all questions are distributed according to the same multivariate normal distribution. The guilty hypothesis assumes that a subject's response magnitudes for noncritical and critical questions are distributed respectively according to each of two multivariate normal distributions. The decision was made by adopting the hypothesis which shows the higher validity for observed magnitudes of responses, where the AIC statistic is taken as a measure for the validity. A feature of this algorithm is to make a decision on the basis of only the record obtained from the judged subject, in contrast with previous methods using the discriminant function which necessitated the past records of many subjects verified to be guilty or innocent.

The computer-based system was evaluated with the data obtained from an experiment. In the experiment, one half of the subjects were guilty of committing a mock theft, and one half were innocent. The computer decisions were 73 % correct on 48 records from innocent subjects and 90 % correct on 48 records from guilty subjects. These computer outcomes were superior to human examiners' judgments.

緒 言

虚偽検出検査の主要な質問法である緊張最高点質問法 (Guilty Knowledge Test) では、質問刺激に対する生理

反応に基づいて、被検者が犯行に関連した情報を知っているかどうか判定される (Lykken, 1959⁽⁶⁾, 1960⁽¹⁷⁾; 鈴木, 1986⁽²⁶⁾; 山岡, 1989⁽³²⁾; Ben-Shakhar & Furedy, 1990⁽⁸⁾)。例えば、絞殺の凶器がネクタイであり、こ

の犯行情報を無実の者が知らない場合には、「ネクタイ」とともに「ストッキング」、「ベルト」、「スカーフ」および「タオル」を質問項目として、「被害者は～で首を絞められましたか」という質問の仕方でも各項目を被検者に提示する。被検者は各質問に対して、通常、「知りません」等の否定の返答を行う。もし、被検者が犯行情報を知っていれば、「ネクタイ」に対してだけ虚偽の返答をすることになり、他の質問刺激とは程度の異なる顕著な生理反応を生起すると考えられる。一方、被検者に犯行情報の知識がなければ、どの質問刺激も同等の意味しか持たないので、質問間で反応量の差が見られないだろう。上例の「ネクタイ」のように犯行情報からなる質問を裁決質問、他の犯行に関連しない質問を非裁決質問と呼ぶ。

こうした緊張最高点質問法に基づく虚偽検出の研究は多岐に渡り、近年に限っても、種々の変数を扱った数多くの研究がある(鈴木, 1986²⁶⁾; Raskin, 1989²²⁾; Ben-Shakhar & Furedy, 1990⁸⁾)。例えば、質問に対する返答の役割は幾つかの実験で取り上げられ、無返答の場合にも裁決質問に対して顕著な反応が生起する傾向が確認されている(Orne, Thackray & Paskewitz, 1972²¹⁾; Horneman & O'Gorman, 1985¹²⁾; Elaad & Ben-Shakhar, 1989⁹⁾)。この結果から、裁決質問に対する反応生起の必要条件は、虚偽の返答ではなく、被検者が非裁決質問を識別できること、言い換えれば、検査時に犯行情報、つまり裁決質問を再認することであるといえよう。こうした裁決質問の再認過程と生理反応との関係については、近年、Gati & Ben-Shakhar (1990¹⁰⁾)によって基礎的研究がなされている。しかしながら、返答をさせる方が無返答条件よりも、裁決質問に対して一層顕著な反応が生じることも知られており(Orne et al., 1972²¹⁾; Horneman & O'Gorman, 1985¹²⁾; Elaad & Ben-Shakhar, 1989⁹⁾)、返答は言わば虚偽検出の増強要因に位置づけられる。返答とともに重要な増強要因は、被検者の虚偽の動機づけであり(Orne et al., 1972²¹⁾; Elaad & Ben-Shakhar, 1989⁹⁾)、動機づけを含めた種々の社会心理的な変数を扱った研究も数多くある(鈴木, 1981²⁵⁾; Waid & Orne, 1981³¹⁾)。以上に例示した諸研究は、虚偽検出の認知的側面の研究としてまとめることができ、応用的には質問構成法や検査手続き等に関連するものであるが、これらの研究領域とともに応用上重要な意義を持つのは、判定に関連する研究である。

現行の虚偽検出検査では、呼吸運動、皮膚電気反応および脈波等の複数種の反応記録を、人間が視察して判定を行う方法が採られている。すなわち、検査者が、裁決質問および非裁決質問の間で複数指標の反応量を比較判

断し、犯行情報の知識の有無が判定される。こうした視察判定は一定の規準に従ってなされるため、検査者間で判定の個人差はあまり見られないものと推論されるが、個人的主観が入り込む余地がないとはいえない。こうした判定の信頼性の問題を回避する方策として、コンピュータによって判定を行うことが考えられる。すなわち、ポリグラフで測定した反応をコンピュータで計測し、計測された反応量の統計解析に基づいて判定を下すわけである。コンピュータ判定の実現により、信頼性の問題の克服および判定の自動化が達成されると同時に、人間の視察を凌ぐ高精度の判定が可能になることも期待できる。そこで、本研究では、コンピュータによる判定システムを提案し、データによってその判定精度を評価する。

提案する判定システムの内容は、(1)反応計測のアルゴリズムと、(2)判定のアルゴリズムに大別できる。前者は、呼吸運動、皮膚電気活動および脈波の各生体現象について、質問刺激に対する反応量を計測するものである。後者は、計測された反応量の統計解析によって意思決定を行なうもので、反応指標が複数種であることから、多変量分布に基づく解析を用いる。この解析手法は、後述するように、過去に虚偽検出検査への適用が試みられた判別分析法(Kircher & Raskin, 1986¹⁴⁾; 足立, 1989¹¹⁾; 足立・鈴木, 1991-a³⁾; 和田・久保, 1991³⁰⁾; 徳田, 1992²⁷⁾, 1993²⁸⁾)とは異なり、我々が新たに導入したものである。

反応計測の方法

虚偽検出検査では、上述したように裁決質問と複数の非裁決質問を質問刺激とし、各質問刺激を15～20秒ほどの間隔をおいて提示する。こうした質問刺激の継時提示を行なう系列(Fig. 1 参照)を複数回繰り返す。各質問刺激に対する反応の程度が判定の基礎データとなる。反応の測定にはLafayette社製のポリグラフを用い、呼吸運動(DC増幅)、皮膚コンダクタンス(または抵抗)反応(時定数4秒)および上腕脈波(DC増幅)が測定されることを前提とする。ポリグラフで測定された反応をA/D変換した後、コンピュータに送り、計測処理を行なう。以下、各チャンネルごとに反応指標の定義および計測方法を記す。なお、以下に記す反応指標は、過去の研究(足立・鈴木, 1991-b⁴⁾)および予備的分析に基づいて、検出精度が比較的高く、かつ、相関が高い指標対が含まれないことを考慮して選定したものである。

1. 呼吸運動

質問刺激に対する返答の直後から返答の10秒後までの

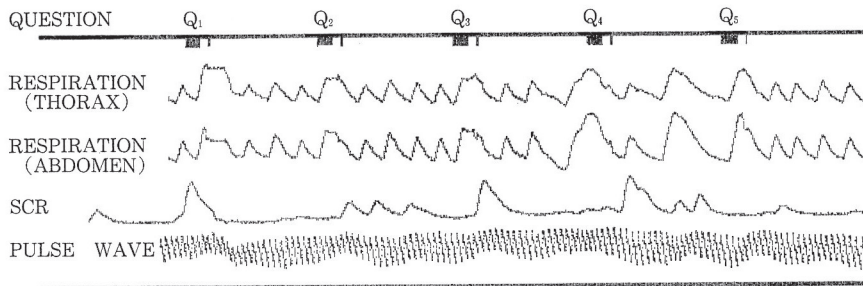


Fig. 1 An example of a digitized record of physiological responses for a series of questions.

間隔を対象とし、呼吸波の垂直方向の移動距離の時間平均をもって呼吸速度とする（足立・鈴木，1991-b⁴⁾；徳田，1993²⁸⁾）。従来から呼吸運動の反応指標として取り上げられてきた振幅、周期および呼吸時間等（中山・水谷，1988²⁰⁾；足立・鈴木，1990²⁾）については、次に列挙する理由から用いなかった。(1)虚偽検出精度は呼吸速度が最も優れており、さらに、呼吸速度には振幅・周期時間等の情報が含まれている（足立・鈴木，1990²⁾）。(2)振幅、周期時間を計測するためには、吸気・呼気完了時点に対応する呼吸波の極大・極小点を同定する必要があるが、不規則な呼吸波については上記の点を自動的に同定することは困難である。

呼吸運動は胸部および腹部について測定されるため、2種の呼吸速度が得られるが、両者を平均し、1つの値にまとめる。胸部・腹部間の呼吸速度は一般に高い相関を示すが、後述する統計解析では相関の高い変数対が含まれない方が望ましいからである。

2. 皮膚コンダクタンス反応

皮膚コンダクタンス反応（以下、SCRと略記）の振幅を反応指標とする。すなわち、質問提示開始0.5秒後から返答後5秒後までの間隔を対象とし、コンダクタンスの最大上昇量を求め、その平方根を反応指標とする。開平変換を行う理由は次のとおりである。無変換の振幅の分布は一般に高い歪度を持つが（Venables & Christie, 1980²⁹⁾），開平変換により分布の左右対称性を増大させると、後述する正規分布を仮定した統計的判定に都合よくなるからである。さらに、開平変換を施した方が虚偽検出精度が高くなること（足立・鈴木；1991-b⁴⁾）も理由の1つである。

3. 脈波

質問提示前3拍の基線の高さを平均し（質問前平均値）、質問提示開始直後から15秒後までの基線の高さの平均値（質問後平均値）から質問前平均値を引いた値を基線変化量（以下、脈波反応量と呼ぶ）とする。

4. 心拍

脈波の頂点間の時間間隔から、1秒ごと（second by second）の心拍率を求め、質問刺激に対する心拍反応量を算出した。算出方法は、Jones, Crowell & Kapuniai (1969³³⁾）の反応変化量検出法を変形したものであり、以下の通りである。まず、心拍率は3次の自己回帰過程（廣松・浪花，1990¹¹⁾）に従うものと仮定し、質問提示開始後15秒間の各秒における心拍率 y_t ($t=1, 2, \dots, 15$) を、それに先立つ3秒間の心拍率 y_{t-3} , y_{t-2} , y_{t-1} に基づいて予測する。すなわち、心拍率 y_t を、

$$\hat{y}_t = \hat{\mu} + \sum_{j=1}^3 a_j (y_{t-j} - \hat{\mu})$$

によって予測する。ただし、 μ は心拍率の平均値、 a_j は自己回帰係数で、1系列を通した心拍率から、予測誤差が独立同一の正規分布に従うという仮定のもとに最尤推定する（北川，1988¹⁵⁾）。そして、観測値 y_t の予測値 $\hat{y}_t$ からの偏差の和

$$\sum_{t=1}^{15} (y_t - \hat{y}_t)$$

をもって心拍反応量とする。

5. 反応の標準化

検査では、生理反応の変動が大きい場合に、記録波形の振幅がオーバースケールする、あるいは、小さくなり過ぎる等のことを防ぐために、系列と系列の間で反応増幅の感度を変えることがある。従って、異なる系列間で反応量の絶対値を比較することはできない。そこで、系列間での反応量の比較を可能にするために標準化を行なう。つまり、反応量を、系列内を通した反応傾向を規準とした相対的な値に変換するわけである。具体的には、以下の方法で標準化を行なう。

呼吸運動、脈波および心拍率については、上述の反応指標の計測・算出の前に、元の波形そのものを標準化する。すなわち、系列の最初の質問提示の直後から最後の質問提示の15秒後までの波形の波高値の標準偏差を求め

(南, 1986¹⁹⁾), この標準偏差で元の波高値を除す。この変換によって, 各系列の波高値の分散, つまり, 波形の垂直方向へのばらつきが等しく 1 となる。SCR は, 開平変換後の振幅について, 足立・鈴木 (1991-b⁴⁾) の方法による系列内標準得点を求める。すなわち, 系列内の各質問刺激に対する振幅の平均を求めた上で, 各振幅値の平均偏差値を算出し, この偏差値を標準偏差で除す。ただし, この標準偏差は, 系列内で求めるのではなく, 全系列の平均偏差をプールした値から求める。

判定の方法

緒言で記したように, 質問刺激は, 検査対象の犯行には関係のない複数の非裁決質問と, 犯行に関連する裁決質問から構成され, これらの質問刺激を提示する系列を繰り返す。裁決および非裁決質問は, 犯行に関与していない者にとって, どれが裁決質問であるか識別できないように選定されるので, 被検者が無実であれば, 全ての質問に対して同程度の反応しか発現しないと考えられる。一方, 被検者に犯行情報の知識があれば, 裁決質問に対して非裁決質問とは程度の異なる反応を発現すると考えられる。従って, 判定は, 各質問に同程度の反応が生じている (陰性) か, あるいは, 裁決質問と非裁決質問の間で有意な反応量の差が見られる (陽性) かを判断することに帰着される。以上の判定を, 数理統計学的に行うことが我々の目的である。

判定の基礎データは, 上述の反応計測のアルゴリズムを通して得られた, 呼吸速度, SCR 振幅, 脈波反応量および心拍反応量の計 4 種の指標である。まず, これら 4 指標の反応量が多変量正規分布 (Anderson, 1984) に従うものと仮定した上で, 被検者が犯行情報を知らない場合と知っている場合のそれぞれについて, 反応の出現の仕方を考える。前者, つまり被検者が陰性の場合には, 全ての質問刺激が同等の意味しか持たないので, 各質問刺激に対する反応量は同一の多変量正規分布に従って生じるものと仮定できる。この仮定を陰性モデルと呼ぶことにする。他方, 後者, つまり被検者が陽性の場合には, 裁決質問が非裁決質問とは異なる意味を持つので, 両種の質問刺激に対する反応量は, それぞれ異なる多変量正規分布に従って生じるものと考えられる。この仮定を陽性モデルと呼ぶことにする。以上より, 判定は被検者から得られた反応データについて, 陰性モデルと陽性モデルのいずれが適合するかを判断することに帰着される。

以上の仮定をより厳密に表わして議論を進めるため, 数式表現を用いる。まず, 質問刺激の数は n 個とし, 各

質問刺激を添え字 j (非裁決質問; $j=1, \dots, n-1$, 裁決質問; $j=n$) で表わす。また, 反復する質問の系列を添え字 i ($i=1, 2, \dots, m$) で表わすと, 系列 i の質問刺激 j に対する複数 (p) 指標の反応量はベクトル

$$\mathbf{x}_{ij} = [x_{ij1}, x_{ij2}, \dots, x_{ijp}]'$$

で表せる。ただし, 本研究で取り上げる指標は上述の 4 種であるので, $p=4$ である。多変量正規分布を $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ($\boldsymbol{\mu}$ は平均ベクトル, $\boldsymbol{\Sigma}$ は共分散行列) で表わすと, 上記の 2 種のモデルは次のように記述できる。

陰性モデル;

$$\mathbf{x}_{ij} \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \quad (i=1, \dots, m; j=1, \dots, n). \quad (1)$$

陽性モデル;

$$\mathbf{x}_{ij} \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \quad (i=1, \dots, m; j=1, \dots, n-1),$$

$$\mathbf{x}_{ij} \sim N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \quad (i=1, \dots, m; j=n).$$

これらの各モデルのいずれが, 観測値 $\mathbf{X} = \{\mathbf{x}_{ij}; i=1, \dots, m; j=1, \dots, n\}$ に妥当するかを比較することによって判定を行う。妥当性の指標は, 尤度に基づいて定義する (芝・渡部, 1984²⁴⁾)。観測値間の独立性を仮定すると, 各モデルの尤度は, 多変量正規分布の確率密度関数 ϕ を用いて次のように表わせる。

陰性モデル;

$$L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0; \mathbf{X}) = \prod_{i=1}^m \prod_{j=1}^n \phi(\mathbf{x}_{ij}; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0). \quad (3)$$

陽性モデル;

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2; \mathbf{X}) = \prod_{i=1}^m \prod_{j=1}^{n-1} \phi(\mathbf{x}_{ij}; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \times \prod_{i=1}^m \phi(\mathbf{x}_{in}; \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2). \quad (4)$$

ただし,

$$\phi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-p/2} |\boldsymbol{\Sigma}|^{-1/2} \times \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right] \quad (5)$$

各モデルのもとで尤度を最大化するパラメータ (母数) の値 (最尤推定値), すなわち, $\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0$ および $\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\Sigma}}_1, \hat{\boldsymbol{\Sigma}}_2$ を求め, 最大尤度 $L(\hat{\boldsymbol{\mu}}_0, \hat{\boldsymbol{\Sigma}}_0; \mathbf{X})$ (陰性モデル) と $L(\hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\mu}}_2, \hat{\boldsymbol{\Sigma}}_1, \hat{\boldsymbol{\Sigma}}_2; \mathbf{X})$ (陽性モデル) の大小比較に基づいて判定を行う。ここで, パラメータ数の節減のため, 陽性モデルの共分散行列について

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 \quad (6)$$

という制約をおく。各パラメータの最尤推定値は以下の式で求める。

陰性モデル;

$$\hat{\boldsymbol{\mu}}_0 = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \mathbf{x}_{ij},$$

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (\mathbf{x}_{ij} - \boldsymbol{\mu}_0)(\mathbf{x}_{ij} - \boldsymbol{\mu}_0)'. \quad (7)$$

陽性モデル；

$$\begin{aligned}\hat{\mu}_1 &= \frac{1}{(m-n)} \sum_{i=1}^m \sum_{j=1}^{n-1} x_{ij}, \\ \hat{\mu}_2 &= \frac{1}{m} \sum_{i=1}^m x_{ij}, \\ \hat{\Sigma}_1 = \hat{\Sigma}_2 &= \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^{n-1} (x_{ij} - \mu_1)(x_{ij} - \mu_1)' \\ &+ \frac{1}{mn} \sum_{i=1}^m (x_{in} - \mu_2)(x_{in} - \mu_2)'\end{aligned}$$

以上の式より、各モデルのもとでの最大尤度を求め、その値が大きい方のモデルを採用する方法が考えられる。しかし、一般に最大尤度は、モデルの良否の測度として偏りをもつことが知られている。すなわち、モデルのパラメータ数が多いほど高くなる偏りを持ち、上述の陰性・陽性モデルでは、2種の分布を仮定する分だけパラメータ数が多い陽性モデルの方が大きい最大尤度を示す傾向にあることが予想される。こうしたパラメータ数の効果を補正したモデルの良否（悪さ）の測度として赤池の情報量規準（Akaike's Information Criterion；AIC）があり、次の式で定義される（Akaike, 1974⁵⁾；赤池, 1976⁶⁾）。

$$AIC = -2 \cdot \ln(\text{最大尤度}) + 2 \cdot K. \quad (9)$$

ただし、 K はモデルに含まれる自由パラメータ数で、陰性および陽性モデルのそれぞれについて、 $K = p + p(p+1)/2$ および $K = 2p + p(p+1)/2$ となる。この AIC を判定に用いる。すなわち、AIC の値が小さいモデルほど良いことになるので、各モデルのもとで AIC を求め、AIC の値の小さい方のモデルをもって判定を下す。例えば、陽性モデルの AIC の値が陰性モデルより小さければ、陽性判定、つまり「被検者は当該犯行情報を知っている」との判定を行う。

判定システムの作成

パーソナル・コンピュータを用いて、自動判定を行うシステムを作成した。これは、Lafayette 社製ポリグラフで測定された生理反応、あるいは、それらのデータレコーダによる磁気記録を、A/D コンバータ（Micro-Science 社製 ADM-1998BPC）によってデジタル化し、パーソナル・コンピュータ上で計測・判定を行うものである。A/D 変換のサンプリング間隔は、呼吸運動および SCR については 100msec、脈波および質問・返答を計時したマーカー信号については 20msec とした。計測・判定は上述のアルゴリズムに従い、それらのプログラムは MS-DOS 上の N88-BASIC によって作成した。

判定結果は、陽性または陰性の 2 カテゴリーを表示するとともに、モデル間の AIC の差が一定の値以下の場合には、判定不能のカテゴリーも加えられるようにした。

データによる評価

コンピュータによる判定システムの精度を、実験データを用いて評価した。さらに、人間による視察判定の精度との比較も行った。

1. データ

被験者：48名の男女大学生（男子21名、女子27名、年齢20～24歳）を被験者とした。

質問刺激：「レターケースの引き出しから、ハンカチの下に置かれた6万円と万年筆を盗む」という模擬窃盗を想定し、「ハンカチ」、「6万円」および「万年筆」を裁決質問とする3種の質問表を作成した。つまり、「現金の上に置いてあった物」に関する質問表を「めがね、セロテープ、はがき、ハンカチ、ピンセット」の5つの質問項目、「盗まれた現金の金額」に関する質問表を「2万円、4万円、6万円、8万円、10万円」の5項目、そして、「盗まれた品物」に関する質問表を「ライター、腕時計、ネクタイピン、万年筆、電卓」の5項目で構成した。

実験手続き：被験者を、犯人群24名、無実群24名に分けた。陽性群の被験者には、上述の模擬窃盗を行なわせた後、「虚偽検出で無実と判定されるように努力するように」と教示した。陰性群の被験者には模擬窃盗を行なわせなかった。以上の手続きは、虚偽検出を行なう実験者が、被験者がいずれの群かを知り得ないように配慮して行なった。この後、上述の3種の質問表の中のいずれか2種の質問表を用いて虚偽検出を行なった。各質問表につき、5つの質問を録音テープの再生によって約20秒間隔で提示して被験者に否定の返答を求める系列を、5回反復した。質問の提示順序は系列ごとに変えた。ただし、裁決質問は系列の最初および最後に位直しないようにした。理由は、系列の最初の質問に対しては、質問内容に関係なく、顕著な反応が生じる傾向があり、系列の最後の質問に対する反応も同様の傾向を持つからである。すなわち、裁決質問を最初または最後に提示すると、被験者が無実であっても裁決質問に顕著な反応を生じる可能性があるため、裁決質問の提示位置を限定した。

反応の測定・記録：呼吸運動（胸部・腹部）、SCR、および脈波を Lafayette 社製ポリグラフによって測定した。呼吸運動は Multi-function ユニットによって直流増幅記録し、SCR は GSG ユニットによって増幅記録

(時定数 4 秒) した。脈波は、被験者の上腕部に巻いたカフに 50mmHg の空気圧を加え、Multi-function ユニットによって直流増幅記録した。以上の生理反応および質問・返答を計時したマーカー信号をデータレコーダ (TEAC 社製 MR 30) で磁気記録した。

2. コンピュータ判定の結果

上記の実験により、計96種 (被験者48名×2 質問表: 無実・犯人の各群につき48種) の反応記録が得られた。各反応記録について、コンピュータによる判定を行なった。すなわち、各質問刺激に対する呼吸速度、SCR 振幅、脈波および心拍反応量を計測し、AIC に基づく判定を行なった。判定結果は Table 1 に示す通りで、無実群について正しく陰性判定できた反応記録の比率は 89.6%、犯人群について正しく陽性判定できた質問表の比率は 72.9% であった。両群を合計すると、正判定率は 81.2%、誤判定率は 18.8% となる。誤判定率を低減させる方策の 1 つは判定不能のカテゴリーを設けることである。一般に、AIC の差が 1 より小さい場合には、モデル間の優劣は付け難いことが知られている (坂元・石黒・北川, 1983²³⁾)。そこで、陰性・陽性モデル間の AIC の

差の絶対値が 1 未満あるいは 0.7 未満未満の場合を判定不能とする 3 件法の判定も行なった。結果は Table 2 に示す通りである。無実・犯人の両群をまとめた誤判定率は、判定不能の基準値を 1 および 0.7 とした場合のそれぞれにつき、12.5% および 14.6% であった。

3. 視察判定の結果

視察による判定の精度を、実験データにより評価した。すなわち、6 名の虚偽検出担当の研究技術者に、上記の反応記録の判定を行なわせた。6 名の判定者を 2 名ずつの 3 群にわけ、さらに、96 種の反応記録を 32 種の記録からなる 3 つの群に分割し、判定者の各群に反応記録の各群を割当てた。判定は、「- (陰性), ± (陽性傾向), + (陽性)」の順位カテゴリーに「判定不能」を加えた 4 カテゴリーに基づいて行なわせた。「±」および「+」をともに陽性判定とみなして集計した判定結果を Table 3 に示した。

4. 各指標ごとのコンピュータ判定

呼吸速度、SCR 振幅、脈波反応量および心拍反応量のそれぞれについて、単一の指標に基づいて、コンピュータ判定を行った。判定手法は、複数指標の場合と全く同じである。すなわち、前述の判定手法の説明では指標数を $p=4$ として議論を進めたが、 $p=1$ として同じ計算をすればよい。Table 4 に無実・犯人の各群の反応記録に関する正判定率および両群をまとめた正判定率を示した。両群をまとめた正判定率は、呼吸速度が最高で、脈波反応量が最低であった。

考 察

提案した計測・判定アルゴリズムに基づくコンピュー

Table 1 Percents of computer decisions on records under innocent and guilty conditions

Condition	Decision	
	Innocent	Guilty
Innocent (n = 48)	89.6%	10.4%
Guilty (n = 48)	27.1%	72.9%

Table 2 Percents of computer decisions with an inconclusive category

(A) The inconclusive decision was made if the AIC difference was less than 1.0.

Condition	Decision		
	Innocent	Guilty	Inconclusive
Innocent (n = 48)	79.2%	4.2%	16.7%
Guilty (n = 48)	20.8%	60.4%	18.8%

(B) The inconclusive decision was made if the AIC difference was less than 0.7.

Condition	Decision		
	Innocent	Guilty	Inconclusive
Innocent (n = 48)	83.3%	6.3%	10.4%
Guilty (n = 48)	22.9%	66.7%	10.4%

Table 3 Percents of six examiners' decisions on records under innocent and guilty conditions

Condition	Decision		
	Innocent	Guilty	Inconclusive
Innocent (n=96)	70.8%	22.9%	6.3%
Guilty (n=96)	29.2%	65.6%	5.2%

Table 4 Percents of correct decisions by computer analyses on each of physiological indices

Index	Condition		Total
	Innocent	Guilty	
Respiration	93.8%	54.2%	74.0%
SCR	77.1%	58.3%	67.7%
Pulse Wave	79.2%	27.1%	53.1%
Heart Rate	83.3%	62.5%	72.9%

タ・システムの判定精度をデータによって評価した。判定結果は、Table 1 および Table 2 に示す通りである。無実群の正判定率が高率で、無実の被験者を陽性判定する誤り (False positive) が低かった。一方、犯人群の正判定率は無実群の正判定率に比べて低かった。以上のコンピュータによる判定の精度が十分に高いかどうかを判断するため、視察による判定の精度との比較を行った。視察判定の結果は Table 3 に要約される。無実および犯人の両群について、誤判定率は、コンピュータ判定より視察判定の方がやや高く、コンピュータ判定の方が優れていると言えよう。ただし、視察判定を行った判定者数は6名と比較的少数であるので、以上の結果が必ずしも一般性の高いものであるとは言えないだろう。

本研究で提案した統計的判定手法の特徴は、判定のためのパラメータを、判定対象の反応記録、つまり個体内の反応データから推定する点である。これとは対照的に、従来、虚偽検出の判定への適用が試みられた判別分析法は、集団の反応データに基づく判定法であるといえる (Kircher & Raskin, 1988¹⁴⁾; 足立, 1989¹⁾; 足立・鈴木, 1991-a³⁾; 和田・久保, 1992³⁰⁾; 徳田, 1992²⁷⁾ 1993²⁸⁾)。すなわち、過去の判別分析によるアプローチは、陰性および陽性が明らかになった複数被験者の反応記録から、陰性または陽性被験者が示す反応の確率分布を推定し、判定対象の反応記録に適用するものであった (McLachlan, 1992¹⁸⁾)。従って、事前に過去の集団データの収集を必要とするが、本研究の手法ではこうしたデータ収集の労力が省ける利点がある。また、一般に生理

反応の出現傾向には個人差が大きい³⁾、複数被験者の反応データを1まとめにして処理する判別分析は、こうした個人差を無視したものであるといえる。一方、本研究の手法は、判定を個体内情報に基づいて行うため、個人差の問題とは無縁である。

以上が、本研究で提案した判定手法の長所である。実験データによる評価結果も上述したように満足できるものであり、本研究の判定手法の有効性が示されたが、短所も考えられる。それは、集団の反応データに基づく手法と比較すると、推定すべきパラメータ数に対してデータ数が相対的に少ないことである。従って、被験者が通常の反応傾向からかけ離れた反応を発現させた場合などは、パラメータの推定値がその影響を受けやすいといえる。幸いにも本研究で提案した手法は、過去の複数被験者のデータからパラメータを推定するように変更したり、あるいは、パラメータに制約を設けて推定すべきパラメータ数を削減する等の改変を容易にできる柔軟性を持つ。こうした種々のパラメータ推定法を検討することが、今後の課題の1つとなろう。

本研究では、計測対象の反応指標として、呼吸速度、SCR 振幅、脈波反応量および心拍反応量の計4種を用いた。これらの各指標に基づく虚偽検出精度は Table 4 に要約される。全般に、単一指標に基づく判定精度は、複数指標を用いた場合 (Table 1, 2) に比べて低く、多変量解析を用いることの意義が窺える。各指標の中では、呼吸速度による判定精度が優れており、これは過去の知見 (足立・鈴木, 1991-a³⁾, 1991-b⁴⁾) と一致する。注目されるのは、過去の研究 (足立・鈴木; 1991-b⁴⁾) で呼吸や皮膚電気反応に劣るとされた心拍反応量が、SCR を凌ぎ、高い判定精度を示した点である。この結果は、心拍反応量の定義に自己回帰モデルによる時系列解析を導入したことによるものと考えられる。判定精度が最も低かった脈波についても、時系列的な定義を導入することによって、精度を向上させ得ると予想される。例えば、徳田 (1992²⁷⁾, 1993²⁸⁾) は、FFT (高速フーリエ変換) を、脈波の計測に導入し、その有効性を確認している。今後、こうした種々の時系列解析的な手

法を含め、新たな反応指標の定義法を検討することにより、コンピュータ判定システムの向上を図ることができると考えられる。

本研究にあたり、ご教示・ご指導をいただきました鈴木昭弘室長（科学警察研究所）に感謝いたします。

文 献

- 1) 足立浩平：虚偽検出のための乱塊配置モデルに基づく判別関数. 科学警察研究所報告法科学編, **42**, 160-167, 1989.
- 2) 足立浩平・鈴木昭弘：虚偽検出における反応指標間の相関構造. 科学警察研究所報告法科学編, **43**, 55-58, 1990.
- 3) 足立浩平・鈴木昭弘：判別関数の緊張最高点質問法への適用. 科学警察研究所報告法科学編, **44**, 139-141, 1991-a.
- 4) 足立浩平・鈴木昭弘：生理指標間の虚偽検出精度の比較. 応用心理学研究, **16**, 33-43, 1991-b.
- 5) Akaike, H.: A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **6**, 716-723, 1974.
- 6) 赤池弘次：情報量基準 AIC とは何か. 数理科学, **153**, 5-11, 1976.
- 7) Anderson, T. W.: An introduction to multivariate statistical analysis, 2nd. ed. Wiley, New York, 1984.
- 8) Ben-Shakhar, G. & Furedy, J. J.: Theories and applications in the detection of deception: A psychophysiological and international perspective. Springer-Verlag, New York, 1990.
- 9) Elaad, E. & Ben-Shakhar, G.: Effects of motivation level and verbal response type on psychophysiological detection in the guilty knowledge test. *Psychophysiology*, **26**, 442-451, 1989.
- 10) Gati, I. & Ben-Shakhar, G.: Novelty and significance in orientation and habituation: A feature-matching approach. *Journal of Experimental Psychology: General*, **119**, 251-263, 1990.
- 11) 廣松 毅・浪花貞夫：経済時系列分析. 朝倉書店, 1990.
- 12) Horneman, C. J. & O'Gorman, J. G.: Detectability in the card test as a function of the subject's verbal responses. *Psychophysiology*, **22**, 330-333, 1985.
- 13) Jones, R. H., Crowell, D. H. & Kapuniai, L. E.: Change detection model for serially correlated data. *Psychological Bulletin*, **71**, 352-358, 1969.
- 14) Kircher, J. C. & Raskin, D. C.: Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology*, **73**, 291-302, 1988.
- 15) 北川源四郎：自己帰帰モデルによる時系列解析. 村上征勝, 田村義保（編著）パソコンによるデータ解析. Pp. 167-180, 朝倉書店, 1988.
- 16) Lykken, D. T.: The GSR in the detection of guilt. *Journal of Applied Psychology*, **43**, 385-388, 1959.
- 17) Lykken, D. T.: The validity of guilty knowledge technique: The effects of faking. *Journal of Applied Psychology*, **44**, 258-262, 1960.
- 18) McLachlan, G. J.: Discriminant analysis and statistical pattern recognition. Wiley, New York, 1992.
- 19) 南 茂夫：科学計測のための波形データ処理. CQ 出版, 1986.
- 20) 中山 誠・水谷充良：裁決質問に対する呼吸変化の検討. 科学警察研究所報告法科学編, **41**, 27-34, 1988.
- 21) Orne, M. T., Thackray, R. I. & Paskewitz, D. A.: On the detection of deception: A model for the study of physiological effects of psychological stimuli. In N. S. Greenfield & R. A. Sternbach (Eds.), *Handbook of psychophysiology*. Pp. 743-785, Holt, New York, 1972.
- 22) Raskin, D. C.: Polygraph Techniques for the detection of deception. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence*. Pp. 247-296, Springer, New York, 1989.
- 23) 坂元慶行・石黒真木夫・北川源四郎：情報量統計学. 共立出版, 1983.
- 24) 芝 祐順・渡部 洋：統計的方法Ⅱ；推測, 増訂版. 培風館, 1984.
- 25) 鈴木昭弘：実験的虚偽検出事態における不安, 達成動機, 皮膚電気活動の関連性. 科学警察研究所報告法科学編, **34**, 153-159, 1981.
- 26) 鈴木昭弘：虚偽検出検査. 新美良純, 鈴木次郎（編）皮膚電気活動, Pp. 192-205, 星和書店, 1986.
- 27) 徳田 豊：ポリグラフ検査の自動判定法に関する検討：線形判別関数による判定. 科学警察研究所報告法科学編, **45**, 87, 1992.
- 28) 徳田 豊：ポリグラフ検査の自動判定システムの実用化の試み. 科学警察研究所報告法科学編, **46**, 22-26, 1993.
- 29) Venables, P. H. & Christie, M. J.: Electrodermal Activity. In I. Martin and P. H. Venables (Eds.), *Techniques in psychophysiology*. Pp. 3-67, Wiley, New York, 1980.
- 30) 和田純一郎・久保孝之：判別分析による呼吸波判定（第2報）：陰性群の判定について. 科学警察研究所報告法科学編, **45**, 86, 1992.

- 31) Waid, W. M. & Orne, M. T. : Cognitive, social, and personality processes in the physiological detection of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, vol. 14, Pp. 61-106, Academic Press, New York, 1981.
- 32) 山岡一信 : ポリグラフ : 逸脱の心理. 令文社, 1989.
-