「保健指導」に関する評価の試み

五十川 直 樹*, 池 邊 淑 子**
坂 本 亘*, 後 藤 昌 司***

A Preliminary Evaluation about Health Guidance

Naoki ISOGAWA*, Toshiko IKEBE**, Wataru SAKAMOTO* and Masashi GOTO***

Since April 2008, Ministry of Health, Labor and Welfare of Japan has carried out Health Checkups and Healthcare Advice with a particular focus on the Metabolic Syndrome which make it obligatory for person aged 40 through 74 to reduce medical expenses and prevent lifestyle-related diseases. However, Kondo (2004) indicates a lack of foundation for health checkup. And we think that it is important to clarify foundation for effect of health checkups and healthcare advice. In this paper, through consistent statistical data analysis, we suggested a method which evaluate directed effect, attempting to figure the doctor's character, based on results of the health checkups carried out in April, 2004. As a result, we confirmed that the doctor classified subjects in the directed group from their body types such as weight and BMI, and in directed effect that it reduced weight and BMI, but it gave increase of TG and decrease of HDL which were likely to develop abnormal lipid metabolism. So, we found that adequate evaluation for effect of healthcare advice after careful consideration about the doctor's character leads to suggestion of scientific foundation for health checkups and healthcare advice.

Key words : power normal distribution, CART, data adaptive discriminant analysis, AUC

キーワード : ベキ正規分布, CART 法, データ適応型判別解析, AUC

1. 序 に 代 え て

生活習慣の変化や高齢者が漸増している事実から「生活習慣病 (成人病)」の有病者および予備軍が増加している。「生活習慣病」とは、喫煙や食生活、飲酒、運動、睡眠などの生活習慣がその発症と進行に深く関

与している疾患の総称であり (表示 1)、高コレステロール血症や糖尿病・高血圧などの生活習慣病の危険因子が重なると脳卒中や虚血性心疾患の発症率が高まる (Kadota, Hozawa, Okamura, Kadowak, Nakamura, Murakami, Hayakawa, Kita, Okayama, Nakamura, Kashiwagi, & Ueshima, 2007; Ninomiya, Kubo, Doi, Yonemoto, Tanizaki, Rahman, Arima, Tsuruyuya, Iida, & Kiyohara, 2007). 厚生労働省はとくに内臓脂肪型肥満に起因する生活習慣病の増加に着目し、日本独自の「内臓脂肪症候群 (メタボリック・シンドローム)」の診断基準を用いて行う「特定健診・保健指導」を 2008 年度 4 月から開始し、40 歳から 74 歳までの対象者に「義務化」した (厚生労働省健康局, 2007). 「特定健診・保健指導」の狙いは「医療費の削減」と生活習慣病の予防措置であるが、その「実施基準」の評価が適切でないこと、および「予防」としての働きはないことが危惧されている (大櫛, 2006 : 2007). また、近藤 (2004) が指摘するように、健康診断に関する根拠を求める批判もあり、健康診断や保健指導での有効性の根拠を明

*大阪大学大学院基礎工学研究科数理学領域
(Division of Mathematical Science, Graduate School of Engineering Science, Osaka University)

連絡先 : 〒 560-8531 大阪府豊中市待兼山町 1-3
E-mail : isogawa@sigmath.es.osaka-u.ac.jp, sakamoto@sigmath.es.osaka-u.ac.jp

**大分県地域成人病検診センター

(Oita Health Check-up Center)

連絡先 : 〒 870-1133 大分県大分市大字宮崎 1415

E-mail : ikebe@oita-h-cuc.jp

***特定非営利活動法人医学統計研究会

(Biostatistical Research Association, NPO)

連絡先 : 〒 560-0085 大阪府豊中市上新田 2-22-10-A411

E-mail : gotoo@bra.or.jp

示することが重要になっている。

本稿では、実際の健康診断データをもとに、一貫した統計的データ解析の過程を用いて被験者像を抽出することで保健指導の影響について評価する方法を示す。この評価方法の例示に用いるのは、2004 年 4 月に大分県で実施された、疾患発生の予防を意図した労働安全衛生法に基づく定期健康診断のデータである。被験者は大分県内の某企業に勤務する男女従業員 1,141 例であり、企業から定期健康診断と健診後の事後指導の介入の業務委託を受けた大分県内の某医療施設にて実施された。そして、健康診断の結果をもとに学会等の診断基準を参考として医療施設にて設定した基準に基づき、「1. 異常なし、2. 放置可、3. 要経過観察、4. 要精査、5. 要治療、6. 治療継続」の 6 段階の指導区分に被験者が分類された。そして、「3. 要経過観察、4. 要精査、5. 要治療」に該当する被験者のうち、一定の基準値に該当する被験者が「治療群」に分類され、診断医により生活指導が必要で改善効果が期待されると判断された被験者が事後指導の対象者として「指導群」に分類された。また、「6. 治療継続」に該当する被験者を「治療継続群」に分類した。指導群、治療群、治療継続群以外のすべての被験者を「非指導・非治療群」に分類した。その後、治療群の被験者には医療機関での受診を推奨し、指導群の被験者には生活改善の指導が行われた。この生活改善の指導は次の手順で実施された。

1. 保健師と管理栄養士が各被験者に健康診断の結果をもとに食習慣、運動習慣、喫煙、飲酒などの生活習慣を振り返らせ、事後指導の動機付けを行う。
2. 各被験者が継続可能な生活改善に関する目標を掲げる。
3. 2. の目標実現に向けて、各被験者に具体的な行動計画についてアドバイスを与える。

その後、治療群および指導群の 2 群については、4 ヶ月後に再び追跡のための健康診断が実施され、臨床検査値に改善がみられる否かが「追証」された。

本稿では、この健康診断のデータを用いて保健指導に対する評価方法を示す。上記の通り、指導群には診断医が健診結果の改善に生活習慣の改善が必要で生活習慣改善の効果が期待できると判断した被験者が含まれる。被験者を指導群に判別した根拠を測定した臨床

検査値から探り、被験者像の具体化を試みることで、その被験者像から生活改善の指導内容について推測し、指導群の被験者に対する指導の有効性について検討する。つまり、健康診断のデータを単なる数字の羅列として捉えるのではなく、データを活用し、適切な解析のもとで被験者像を抽出することで、新たな知見を得ることが本解析の目的である。

2 節では、本稿での対象データの解析目標を整理し、その動機を明瞭にする。3 節では、2 節の目標に基づいて抽出された被験者像のうえで、保健指導の影響を調べる。4 節では、3 節の解析結果を保証するために本稿で用いたモデルの適切性を調べるための統計的診断を行い、結果の安定性などを吟味する。

2. 健康診断データの解析

本稿での解析に用いる臨床検査値は、体重、BMI、収縮期血圧、拡張期血圧、総コレステロール (TC)、中性脂肪 (TG)、HDL コレステロール (HDL) である。被験者は 1,141 例(男性 543 名、女性 598 名)であった。治療群の規定は一定の基準値 (収縮期血圧: 160 以上

表示 1. 生活習慣病に関する連関要因図

表 1. 被験者プロフィール

性別	年齢	非指導・ 非治療群	指導群*	治療群	治療 継続群
男性	10～29	76	3 (3)	2	0
	30～39	119	26 (14)	13	1
	40～49	81	45 (34)	17	2
	50～	82	42 (38)	22	6
	男性計	358	116 (89)	54	9
女性	10～29	149	3 (2)	1	0
	30～39	100	3 (3)	2	0
	40～49	96	14 (12)	3	1
	50～	144	32 (28)	33	8
	女性計	489	52 (45)	39	9

(*) 指導群のカッコ内の数字は、4ヵ月後に再び健康診断を受診した被験者の数を表す

かつ拡張期血圧：100 以上，TC：90 未満または 260 以上，TG：250 以上，HDL：25 以下のいずれかに該当）に基づき，指導群の規定は診断医により生活指導が必要で改善効果が期待されると判断された被験者である。指導群，治療群，治療継続群以外の被験者はすべて非指導・非治療群に含まれる。ただし，全被験者の内の 11 例には TC，TG または HDL の臨床検査値に欠測値があったため，本稿の解析からはこれらに該当する被験者を除いた。また，非指導・非治療群から 1 例，指導群から 3 例の被験者には治療群の基準値を上回る収縮期血圧と拡張期血圧が記入されていたため，これらも除いた。非指導・非治療群，指導群，治療群，治療継続群の性，年齢に基づく被験者のプロフィールを表 1 に示す。指導群のカッコ内の数字は，4ヵ月後に再び追跡のための健康診断を受診した被験者の数を表す。ちなみに治療群の被験者の内，血圧が基準値を上回る被験者は 2 例，TC が基準値を上回る被験者は 56 例，TG が基準値を上回る被験者は 42 例であり，HDL が基準値を下回る被験者はいなかった。

本稿での解析目標を以下に列挙する。

目標 1. 診断医による非指導・非治療群と指導群，治療群の分類の根拠を探索し，被験者像で分類の結果を具体的に吟味することで指導内容について推測する。

目標 2. 指導群への指導効果を評価し，指導の有効性について検討する。

目標 3. 解析結果の適切性を吟味するためにデータと本稿で用いたモデルについて統計的診断を行う。

これらの目標を達成するために丸尾・白旗・後藤・駒澤（2008）に倣い，解析全体の流れにおいて論理の整合性が保たれることに留意する。1 節で既述した通り，健康診断の結果をもとに分けられた 6 段階の指導区分のうち，1～5. に該当する被験者が「非指導・非治療群」，「指導群」，「治療群」に分類される。本稿では指導群」に分類された被験者像の抽出に焦点をあて，診断医がどのような基準で「非指導・非治療群」と「指導群」に分類したのかということに注目するため，すでに何らかの治療を行っている「治療継続群」は本解析の趣旨とは異なるので本解析の対象からは除外した。ただし，「非指導・非治療群」と「指導群」の分類に寄与する臨床検査値を探索する際に，「治療群」は打ち切り項を考慮したり，分類回帰樹木（Classification and Regression Tree：CART）法（Breiman, Friedman, Olshen, & Stone, 1984；杉本・下川・後藤，2005）において分岐と基準値との乖離から被験者像を推測するときに用いるため，本解析に使用する。

3. 統計的データ解析の過程

一般に臨床検査値が正值であることに注意して，健康診断で測定された諸種の臨床検査値がベキ正規分布に従うと想定する。ベキ正規分布はデータ適応型分布であり，データによって分布の形状が変化するためにデータに外れ値が存在するかどうかを確認することは非常に重要である。したがって，臨床検査値における外れ値の診断（標本診断）をデータ適応型確率プロット（下川・後藤，2002）で行う。データ適応型確率プロットを図示することでデータが単なる分布の裾領域に入るか，あるいは外れ値であるかを視覚的に評価できる。ただし，視覚的に外れ値の判断をしかねるデータについては米国臨床検査標準化協会（National Committee for Laboratory Standard：NCCLS，現 CLSI[Clinical and Laboratory Standards Institute]）のガイドライン（Sasse, Dumas, Miller, D'Orazio, Eckfeldt, Evans, Graham, Myers, Parsons, & Stanton, 2000）で推奨されている評価方法である Dixon 比（[極端な測定値と 2 番目に大きい（小さい）測定値との絶対偏差]/[極端な値を含む全測定値の幅]）に基づく方法（Dixon，

1953) により計量的に判断する。この Dixon 比が $1/3$ 以上のとき、極端な測定値を除去する。外れ値の診断を行った後、説明変数の交互作用や非線形効果を捉え、データに影響を及ぼす説明変数を最適に選択する CART 法とデータが多変量ベキ正規分布に従う場合に高い正判別率を示すことが知られているデータ適応型判別解析 (畠中・井上・後藤, 1981; Seo, Shimokawa, Daimon, & Goto, 2002; 下川・後藤, 2004) により、非指導・非治療群と指導群・治療群の分類に寄与した臨床検査値を探索する。さらに、諸種の臨床検査値の指導前後の値が 2 変量ベキ正規分布 (Goto & Hamasaki, 2002) に従うと想定したうえで、分布形状と指導前後の変動の大きさを求め、指導群に分類した被験者に行った生活改善の指導効果を測る。

ベキ正規分布、データ適応型判別解析の方法の概略は補遺に与える。

3.1. 非指導・非治療群と指導群の分類に寄与した臨床検査値の探索

外れ値の診断: 非指導・非治療群と指導群の男女別のデータ適応型確率プロットを図 1, 図 2 で示した。図 1, 図 2 では従来からの外れ値の評価方法である平均値 ± 3 SD 以上の観測値を含む臨床検査値を表示させている。これらの図より、ほとんどの観測値が曲線に沿って点在していることから各臨床検査値がベキ正規分布に従っていることがわかる。女性 (非指導・非治療群) の HDL の最大値に変換曲線からの乖離がみられたが、この観測値に基づく Dixon 比は $0.179 < 1/3$ となり、Dixon 比に基づく方法では外れ値という診断を下すことができない。ちなみにいずれの観測値も Dixon 比に基づく方法では外れ値となる観測値は存在しなかった。図 1, 図 2 で記していない臨床検査値についてもデータ適応型確率プロットから外れ値が存在しないことを確認した。以上の結果から、いずれの観測値も除去せずに以降の解析で用いる。

CART 法: CART 法を用いて、体重、BMI、収縮期血圧、拡張期血圧、TC、TG、HDL の臨床検査値、年齢、性別 (2 値) を説明変数として、非指導・非治療群と指導群への分岐パターンを調べた (図 3, 表 2)。最適樹木の選択には交差確認法 (1SE ルール) を用いた。その結果、BMI、体重、年齢の順に分岐された。このとき、非指導・非治療群と指導群の分類における誤分類率は 13.69% である。この結果から分類には体型の特徴が効いていることがわかる。ただし、指導群

と治療群は一定の基準値によって判別されているため、実際には指導群の臨床検査値には基準値による打ち切りが生じていると考えられる。この影響を測るため、指導群と治療群を併合し、非指導・非治療群と指導・治療群への分岐パターンを CART 法を用いて調べた (図 4)。その結果、BMI のほかに主に被験者を治療群に識別した臨床検査値である TC または TG での分岐がみられた。TC は治療群の規定 (260 以上) にはほぼ一致した分類結果になっていることに對し、TG は治療群の規定への基準値 (250 以上) よりも低い基準値で指導・指導群に分類されたことから、医師は指導群への識別の際に TG の数値も若干考慮に入っていたのかもしれない。

データ適応型判別解析: 上記の分類結果の確認・診断と判別予測のために、データ適応型判別解析を行った。体重、BMI、収縮期血圧、拡張期血圧、TC、TG、HDL について、非指導・非治療群と指導群の分類に寄与した項目を探索、評価した。CART 法による結果では性別の分岐が現れなかったが、TC、TG、HDL に性別の影響が存在することを示した丸尾ら (2008) の解析結果に倣い、男女別にデータ適応型判別解析を行った。上記の 7 個の臨床検査値による判別の正分類割合は男性 76.0% (非指導・非治療群の正分類割合: 74.3%, 指導群の正分類割合: 81.0%), 女性 83.6% (非指導・非治療群の正分類割合: 83.0%, 指導群の正分類割合: 88.5%) であった。また、受信者操作特性曲線 (Receiver Operating Characteristic Curve: ROC 曲線) の AUC が最も高い変数の組み合わせを選択すると、男性では BMI、収縮期血圧、TC、女性では BMI、拡張期血圧、TC であった (図 5)。この選択された変数を用いて非指導・非治療群と指導群の分類を行った結果、男性 76.0% (非指導・非治療群の正分類割合: 73.5%, 指導群の正分類割合: 83.6%), 女性 81.2% (非指導・非治療群の正分類割合: 80.2%, 指導群の正分類割合: 90.4%) であった。これより、男性と比較して女性の方に明確な判断基準があったのかもしれない。ここでも指導群の臨床検査値に対する打ち切りの影響を測るため、指導群と治療群を併合し、非指導・非治療群と指導・治療群に対してデータ適応型判別解析を行った。その結果、7 個の臨床検査値による判別の正分類割合は男性 80.7% (非指導・非治療群の正分類割合: 79.9%, 指導・治療群の正分類割合: 82.4%), 女性 85.3% (非指導・非治療群の正分類割合: 83.6%, 指導・治療群の正分類割合: 83.6%) であった。

非指導・非治療群(男性)

指導群(男性)

図 1. 外れ値の診断: 男性

非指導・非治療群(女性)

図2. 外れ値の診断: 女性

表2. CART法(非指導・非治療群と指導群の分類)で抽出された被験者像

終結ふし	BMI		体重		年齢		非指導・ 非治療群	指導群
	<24.35	≥24.35	<62	≥62	<30.5	≥30.5		
非指導・非治療1	○						712	54
非指導・非治療2		○	○				36	4
非指導・非治療3		○		○	○		22	4
指導		○		○		○	77	106

図3. 非指導・非治療群と指導群の分類 (CART 法)

図4. 非指導・非治療群と指導・治療群の分類 (CART 法)

合：94.5%)となり、非指導・非治療群と指導群の正分類割合よりも高かった。また、男性はBMI、収縮期血圧、TC、女性はBMI、拡張期血圧、TCを用いて非指導・非治療群と指導・治療群の分類を行った結果、男性76.7%（非指導・非治療群の正分類割合：74.9%，指導・治療群の正分類割合：80.6%），女性84.1%（非指導・非治療群の正分類割合：83.0%，指導・治療群の正分類割合：90.1%）となり、再び非指導・非治療群と指導群の正分類割合よりも高かった。

CART法とデータ適応型判別解析による分類の結果から、指導群に分類する被験者に対して主に体型および血圧に関する指導が行われたと推察できる。

3.2. 臨床検査値の指導前後での分布形状と変動の大きさ

3.1節で抽出した被験者像から推察した指導群の被験者に対する指導内容を念頭にその指導効果を測るために、諸種の臨床検査値の指導前と指導後の観測値が2変量ベキ正規分布に従うと想定する。表3に指導前と指導後の観測値のベキ・パラメータの推定値 $\hat{\lambda}_{前}$ 、 $\hat{\lambda}_{後}$ を示す。この表から、一般に考えられているように、ほとんどの臨床検査値の指導前と指導後の観測値が従う分布形状はほとんど変わらなかったが、男性の拡張期血圧、TC、女性のHDLで $\hat{\lambda}_{前}$ と $\hat{\lambda}_{後}$ の差が大きく、指導前と指導後の観測値が異なる分布に従うことが示唆された。

指導前後で分布形状がほとんど変わらなかった臨床検査値に対して、指導効果を測るために、指導前と指導後の観測値に対するベキ変換後の正規性を利用して、対応のある2標本のt検定（片側：有意水準0.05）を行った。表3の結果から、指導前と指導後の観測値が男性の体重、収縮期血圧、そして女性の収縮期血圧では平方根変換（ $\hat{\lambda}_{前}=\hat{\lambda}_{後}=0.5$ ）、男性のBMI、TGでは対数変換（ $\hat{\lambda}_{前}=\hat{\lambda}_{後}=0$ ）、男性のHDL、女性の体重、BMI、TGでは逆平方根変換（ $\hat{\lambda}_{前}=\hat{\lambda}_{後}=-0.5$ ）、女性の拡張期血圧、TCでは1.5乗変換（ $\hat{\lambda}_{前}=\hat{\lambda}_{後}=1.5$ ）の2変量ベキ正規分布に従うと想定した。 $\mu_{前}$ 、 $\mu_{後}$ をそれぞれ指導前、指導後のベキ変換後の観測値の平均とし、帰無仮説 $H_0: \mu_{前}=\mu_{後}$ に対して、対数変換・平方根変換・1.5乗変換の2変量ベキ正規分布ならば対立仮説 $H_1: \mu_{前}>\mu_{後}$ 、逆平方根変換の2変量ベキ正規分布ならば対立仮説 $H_1: \mu_{前}<\mu_{後}$ （ $H_1: \mu_{前}>\mu_{後}$ [男性のHDLのみ]）の対応のある2標本のt検定（片側）を行った。その結果、男性は体重、BMI、収縮期血圧、女性は体重、BMIで有意に減少していた。しかし、逆方向の対立仮説 H_1 （たとえば、逆平方根変換の2変量ベキ正規分布ならば対立仮説 $H_1: \mu_{前}>\mu_{後}$ ）に対して検定した結果、男性のTG、女性の収縮期血圧とTGでは有意な増加が認められ、男性のHDLでは有

表3. 指導前と指導後のベキ・パラメータの推定値：指導群

	体重		BMI		収縮期血圧		拡張期血圧		TC		TG		HDL	
	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$	$\hat{\lambda}_{前}$	$\hat{\lambda}_{後}$
男性	0.31	0.29	0.23	0.15	0.49	0.81	1.11	2.04	1.54	0.65	0.04	-0.11	-0.51	-0.16
女性	-0.21	-0.41	-0.19	-0.39	0.59	0.42	1.51	1.71	1.46	1.50	-0.35	-0.22	0.70	0.17

図5. 非指導・非治療群と指導群の分類：AUC (1. 体重, 2. BMI, 3. 収縮期血圧, 4. 拡張期血圧, 5. TC, 6. TG, 7. HDL)

意に減少していた。また、指導前と指導後の観測値が従う分布が異なることが示唆された男性の拡張期血圧とTC、女性のHDLのスライディング・スクエア・プロットを図6に示す。この図から、とくに女性のHDLが指導後に減少していることが伺える。スライディング・スクエア・プロットは指導前後差の全体像を見るために役だが、ちなみにノンパラメトリック検定法であるWilcoxon符号つき順位検定(有意水準: 0.05)を用いて中央値が0であるかを調べた結果、男性の拡張期血圧と女性のHDLが有意になった。

3.3. 考察

3.1節から、明確な規定のなかった非指導・非治療群と指導群の分類には主に体重、BMIといった体型の特徴が大きく関与していた。さらに、図5のAUCの値から、収縮期血圧、拡張期血圧、TCも非指導・非治療群と指導群の分類に寄与していると考えられるため、指導群に分類した被験者には主に体型と血圧に関する指導が行われたことが推測される。また、データ適応型判別解析における正分類割合の高さから男性と比較して女性の方に明確な判断基準があったのかもしれない。これらの結果をふまえたうえで、3.2節で指導効果を測った結果、男性の体重、BMI、収縮期血圧、女性の体重、BMIの臨床検査値は有意に減少し、指導による改善効果を示唆したが、男性のTG、

HDL、女性の収縮期血圧、TG、HDLの臨床検査値は悪化していた。これらの結果のみで指導の有効性を判断することはできないが、男女ともに分類に大きく関与していた体重、BMI(体型の特徴)の臨床検査値は減少したにも関わらず、脂質代謝異常につながるTGとHDLの臨床検査値が悪化したことに注意を向けなければならない。以上のことから、被験者像の具体化が指導効果の結果を解釈することに役立つことがわかった。

4. 統計的診断

3.1節では、臨床検査値が従う分布としてベキ正規分布を想定し、非指導・非治療群と指導群の分類に寄与した臨床検査値の探索を行った。ただし、データ解析により得られた結果(知見)を保証するためには統計的診断(「標本診断」、「構造診断」)を行うことが必要である。標本診断(外れ値や影響観測値の診断)はすでに行ったため、ここではモデル(ベキ正規分布)に対する構造診断を行う。図1、図2のデータ適応型確率プロットから諸種の臨床検査値に対するベキ正規分布の整合性を視覚的に判断することができる。図内の変換曲線と観測値の乖離の小ささからどの臨床検査値もベキ正規分布によく適合していることがわかる。さらに、ベキ正規分布の中央値(50%点)を以下の

男性

女性

図 6. スライディング・スクエア・プロット：指導群

手順で点検した。

- 点検 1. 各臨床検査値の中央値をベキ正規分布から推定する。
- 点検 2. 各臨床検査値のブートストラップ標本から中央値を求める作業を 1,000 回繰り返す。
- 点検 3. 点検 2. で求めた中央値のヒストグラムを作成し、点検 1. のベキ正規分布に基づく中央値と比較する。

この結果を図 7 と図 8 に示す。図内の中央の目盛りは分布から推定された中央値である。非指導・非治療群

について、男性の収縮期血圧、HDL、女性の収縮期血圧、拡張期血圧以外の臨床検査値は分布から推定した中央値をほぼ軸にブートストラップ標本の中央値が集合していることから、潜在基礎分布に想定したベキ正規分布が（中央値に関して）適切であると判断できる。ただし、男性の収縮期血圧、HDL、女性の収縮期血圧、拡張期血圧の分布から推定した中央値はブートストラップ標本の中央値のヒストグラム内では右裾部分に位置するため、ベキ正規分布の適切性が疑われるかもしれないが、ブートストラップ標本の中央値の変動が小さいため、実際の現象（データ）とモデル（ベキ正規分布）が大きく乖離していないと考えられる。

図7. 中央値プロット：非指導・非治療群 [中央の目盛りは分布から推定した中央値を示す]

指導群について、ほとんどの臨床検査値は分布から推定した中央値をほぼ軸にブートストラップ標本の中央値が集合しているが、被験者数が少ないために非指導・

非治療群と比べてブートストラップ標本の中央値の変動が大きい。

図8. 中央値プロット：指導群

5. 結びに代えて

本稿では、保健指導に対する評価方法を提示するため、健康診断の結果をもとに、被験者像を具体化し、指導群の被験者に対する指導内容を推測することで、その指導効果に関する評価および考察を行った。健康診断データという「冷たい数字の集まり」(後藤、

1986)を一貫性および論理の整合性を意識したデータ解析過程に通すことで、保健指導に関する評価および考察という生産的知見を獲得することができた。本稿で示した評価方法は他の分類区分でも明確な規定が無い群の特徴を表現したいといった場面に適用が可能である。しかし、今回の「指導群」のように上限または下限に打ち切りが発生する場合にはその影響にも注意

を向けなければならない。保健指導を通じた生活改善の指導効果、疾病発生の予防効果を明確にするためには、本稿で例示したように、統計的診断を含む丁寧な統計的データ解析の過程を経て得られた被験者像に基づいて保健指導の影響を詳細に検討することが重要になる。ただし、4節の結果が示すように小標本に想定するモデルの選択は慎重にすべきであり、本稿の内容のみでモデルの適切性や結果の安定性が十分に検討されたとは言い難いため、実際の適用場面ではさまざまな視点からアプローチすることが必要になる。

謝 辞

本稿の作成にあたり、データの提供にご協力いただいた方々と有益なご指摘と貴重なご意見をいただきました審査員の先生方および編集委員の先生方に心より感謝の意を表します。

参 考 文 献

- Box, G.E.P. & Cox, D.R. (1964). An analysis of transformations (with discussion). *J. Roy. Statist. Soc.*, **B26**(2), 211-246.
- Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification And Regression Trees*. Wadsworth.
- Dixon, W.J. (1953). Processing data for outliers. *Biometrics*, **9**, 74-89.
- 後藤昌司 (1986). 統計的データ解析の過程. 行動計量学, **13**(2), 48-63.
- Goto, M. & Hamasaki, T. (2002). The bivariate power-normal distribution. *Bulletin of Informatics and Cybernetics*, **34**(1), 29-49.
- Goto, M., Matsubara, Y., & Tsuchiya, Y. (1983). Power-normal distribution and its applications. *Rep. Stat. Appl. Res.*, JUSE, **30**, 8-28.
- Goto, M., Uesaka, H., & Inoue, T. (1979). *Some linear models for power transformed data*. Invited paper at the 10th International Biometric Conference. August, 6-10 (Res. Rep. NO.93, Res. Instit. Fund. Infor. Sc., Kyusyu University).
- 畠中駿逸・井上俊昭・後藤昌司 (1981). ベキ正規分布上の判別. 日本行動計量学会第9回大会発表論文, 113-114.
- Kadota, A., Hozawa, A., Okamura, T., Kadowak, T., Nakamura, K., Murakami, Y., Hayakawa, T., Kita, Y., Okayama, A., Nakamura, Y., Kashiwagi, A., & Ueshima, H. for the NIPPON DATA Research Group. (2007). Relationship between metabolic risk factor clustering and cardiovascular mortality stratified by high blood glucose and obesity. *Diabetes Care*, **30** (6), 1533-1538.
- 近藤 誠 (2004). 成人病の真実. 文芸春秋.
- 厚生労働省健康局 (2007). 標準的な健診・保健指導プログラム (確定版): 概要. 厚生労働省健康局. <http://www.mhlw.go.jp/bunya/kenkou/seikatsu/pdf/01.pdf>
- 丸尾和司・白旗慎吾・後藤昌司・駒澤 勉 (2008). 臨床検査値の参照範囲: その根拠と評価. 行動計量学, **35**(1), 73-89.
- Ninomiya, T., Kubo, M., Doi, Y., Yonemoto, K., Tanizaki, Y., Rahman, M., Arima, H., Tsuryuya, K., Iida, M., & Kiyohara, Y. (2007). Impact of metabolic syndrome on the development of cardiovascular disease in a general Japanese population: the Hisayama Study. *Stroke*, **38**, 2063-2069.
- 大櫛陽一 (2006). 検査値と病気. 間違いだらけの診断基準. 太田出版.
- 大櫛陽一 (2007). 読売新聞 (07/10/9 付).
- Sasse, E.A., Doumas, B.T., Miller, W.G., D'Orazio, P., Eckfeldt, J.H., Evans, S.A., Graham, G.A., Myers, G.L., Parsons, P.J., & Stanton, N.V. (2000). *How to Define and Determine Reference Intervals in the Clinical Laboratory: Approved Guideline-Second Edition*. NCCLS.
- Seo, H., Shimokawa, T., Daimon, T., & Goto, M. (2002). *Data-adaptive discrimination based on multivariate power normal distribution*. Contributed paper at the 5th of The International Association of Statistical Computing, ARS, 114-115.
- 下川敏雄・後藤昌司 (2004). データ適応型判別解析とその診断. 計算機統計学, **17**, 87-108.
- 下川敏雄・後藤昌司 (2002). データ適応型確率プロットとその応用. 行動計量学, **29**, 103-119.
- 杉本知之・下川敏雄・後藤昌司 (2005). 樹木構造接近法と最近の発展. 計算機統計学, **18**(2), 123-164.

(2009年7月27日受付, 2010年7月13日最終修正)

補遺: 臨床評価に用いた統計的方法

ベキ正規分布

諸種の臨床検査値の値が従うと想定したベキ正規分布は以下のように与えられる。正值変数 X のベキ変換は

$$Z = \begin{cases} (X^\lambda - 1)/\lambda, & \lambda \neq 0 \\ \log X, & \lambda = 0 \end{cases}$$

と定義される (Box & Cox, 1964)。変換後の $\{Z\}$ の正規性を目標として、モデルの線形性と分散一定性を仮定したときの原観測変数 (X) 尺度上の分布としてベ

キ正規分布が提案されている (Goto, Uesaka, & Inoue, 1979; Goto, Matsubara & Tsuchiya, 1983). その確率密度関数は

$$f_{\text{PN}}(x; \lambda, \mu, \sigma) = \begin{cases} x^{\lambda-1} \phi\{(x^\lambda - \mu)/\sigma\} / A(\lambda, \mu, \sigma), & x > 0 \\ 0, & x \leq 0 \end{cases}$$

である. ここに, $\phi(\cdot)$ は標準正規分布の確率密度関数であり, $A(\lambda, \mu, \sigma)$ は確率比例定数項

$$A(\lambda, \mu, \sigma) = \begin{cases} \Phi\{-(\lambda\mu+1)/\lambda\sigma\}, & \lambda < 0 \\ 1, & \lambda = 0 \\ \Phi\{(\lambda\mu+1)/\lambda\sigma\}, & \lambda > 0 \end{cases}$$

である. ここに, $\Phi(\cdot)$ は標準正規分布の累積分布関数である. パラメータ λ, μ, σ はそれぞれ分布の形状, 位置, 尺度を表すパラメータであり, 原観測変数 X に応じて λ (ベキ・パラメータと呼ぶ) を変化させることで多様な分布に適合する. 一般に $A(\lambda, \mu, \sigma)=1$ を考慮に入れてパラメータを推定することは困難であるため, 本稿では, Box & Cox (1964) の推定方式に倣い, $A(\lambda, \mu, \sigma)=1$ とし, Newton-Raphson 法に基づきパラメータの推定を行う.

p 変量ベキ正規分布 $f_{\text{MPN}}(\mathbf{x} | \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ は, 非負の p 変量観測値 $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$ に対して変量ごとにベキ変換を施して得られる $\mathbf{z} = (z_1, z_2, \dots, z_p)$ が近似的に p 変量正規分布に従うとき, 変換前の $\mathbf{x}$ を規定する分布として定義される. ここに, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)^T$ は, $p \times 1$ のベキ・パラメータ・ベクトルであり, $\boldsymbol{\mu}$ と $\boldsymbol{\Sigma}$ は, それぞれ, 変換後の $\mathbf{z}$ が近似的に p 変量正規分布に従うときの平均ベクトルと分散共分散行列である. 単一変量の場合と同じく, 確率比例定数項 $A(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ を考慮に入れてパラメータを推定することは困難であるため, $A(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\Sigma})=1$ と仮定して, Newton-Raphson 法に基づきパラメータの推定を行う. p 変量ベキ正規分布については Goto et al. (1979), 畠中他 (1981), 下川・後藤 (2004) が詳しい.

データ適応型判別解析

非指導・非治療群と指導群の分類に寄与した臨床検査値の探索に用いたデータ適応型判別解析 (畠中他, 1981; Seo et al., 2002; 下川・後藤, 2004) について記述する. 2 群の p 変量ベキ正規母集団 Π_l ($l=1$: 非指導・非治療群, $l=2$: 指導群) から標本抽出された n_l 個の非負の p 変量観測値 $\{x_{li}\}_{i=1}^{n_l}$ が用意されているとする. $\boldsymbol{\lambda}_l, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l$ が既知であるとき, データ適応型の判別関数は

$$\begin{aligned} g(\mathbf{x}) &= \log f_{\text{MPN}}(\mathbf{x} | \boldsymbol{\lambda}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\mu}_1) - \log f_{\text{MPN}}(\mathbf{x} | \boldsymbol{\lambda}_2, \boldsymbol{\Sigma}_2, \boldsymbol{\mu}_2) \\ &= \frac{1}{2} \left\{ (\mathbf{z}^{(\lambda_2)} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_2^{-1} (\mathbf{z}^{(\lambda_2)} - \boldsymbol{\mu}_2) \right. \\ &\quad \left. - (\mathbf{z}^{(\lambda_1)} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{z}^{(\lambda_1)} - \boldsymbol{\mu}_1) \right\} \\ &\quad + \sum_{i=1}^p (\lambda_{1i} - \lambda_{2i}) \frac{1}{2} \log x_i - \log \left| \frac{\boldsymbol{\Sigma}_2}{\boldsymbol{\Sigma}_1} \right| \end{aligned}$$

で与えられる. 新規に得られた $\mathbf{x}$ に対して, $g(\mathbf{x}) > 0$ ならば Π_1 , $g(\mathbf{x}) < 0$ ならば Π_2 に判別する. 通常, $\boldsymbol{\lambda}_l, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l$ は未知であるため, 最尤推定量で置き換えられる.

また, p 個の説明変数の中の p_r 個の説明変数が判別に寄与する効果は ROC 曲線の曲線下面積 (Area Under Curve: AUC) で評価される. グループ変数 $l(\Pi_l; l=1, 2)$ と説明変数ベクトル $\mathbf{x}$ が与えられたとき, 感度 $F_{\text{TP}}(g, u)$ および特異度 $F_{\text{TN}}(g, u)$ は, それぞれ

$$\begin{aligned} F_{\text{TP}}(g, u) &= \Pr(g(\mathbf{x}) > u | l=1), \\ F_{\text{TN}}(g, u) &= \Pr(g(\mathbf{x}) < u | l=2) \end{aligned}$$

で与えられる. ROC 曲線は, 任意の $u (-\infty < u < \infty)$ に対して, $(F_{\text{TP}}(g, u), 1 - F_{\text{TN}}(g, u))$ をプロットすることで得られる. さらに, AUC は

$$AUC = \int_{u=-\infty}^{\infty} F_{\text{TP}}(g, u) dF_{\text{FP}}(g, u)$$

により算出される. ここに, $F_{\text{FP}}(g, u) = 1 - F_{\text{TN}}(g, u)$ である.