

生理的年齡予測の重回帰分析

吉川, 和利
九州大学健康科学センター

<https://doi.org/10.15017/437>

出版情報 : 健康科学. 7, pp.1-9, 1985-03. Institute of Health Science, Kyushu University
バージョン :
権利関係 :

生理的年齢予測の重回帰分析

吉 川 和 利

A Multiple Regression Model for Physiological Age —On the Case of Females—

Kazutoshi KIKKAWA

The purpose of this study was to establish the multiple regression equation to predict females' physiological age.

Data for this study were collected during the course of survey of various indices on the degree of health and fitness and life-style of healthy women who aged 23 through 59 years old.

Collected information included the followings: systolic blood pressure (=SYSBP), plasma creatinine (=CREAT), plasma albumin (=ALBUM), sum of skinfold thickness of upperarm and subscapulars (=FAT), grip strength (=GRIP), back strength (=BACK), standing trunk flexion (=FLEX), vital capacity (=VITAL), one-foot blind balance (=BALANCE), maximal aerobic power (=MAP, by Margaria's method), and chronological age (=AGE; dependent variables).

Forward stepwise regression analysis examined the reliability of the physiological and internal medicine and anthropometric measurements as the indicator of age.

At the each step, the statistics of degree of fitness, that is, Akaike's *AIC*, Mallows' *C_p* and *R** (adjusted for d.f. *R*) adding to multiple correlation coefficient *R* and *SEE* (standard error of estimates) were calculated.

The results obtained were as follows.

1) In terms of following variables order, all variables were entered into equation; SYSBP, BALANCE, VITAL, FAT, ALBUM, CREAT, FLEX, BACK, GRIP. The multiple correlation coefficient *R*, and standard error of estimates (*SEE*) were 0.738, 6.1 (age of years), respectively.

2) Each equation obtained at each step, which included *m* independent variables at *m* step, in general, were evaluated by several statistics such as Akaike's *AIC*, Mallow's *C_p*, *R** and *R*. According to *C_p*, the equation III which had SYSBP, BALANCE and VITAL was evaluated as the best equation. According to *AIC*, the equation IV which was added FAT to preceding equation (III) was evaluated as best.

3) By the forward stepwise regression analysis, of which stopping rule for variable selection was *F*-statistics=2.0, equation IV was selected as the best equation. Equation IV was described as follows and *R*=0.721, *SEE*=5.9 (age of years), respectively. $\hat{Y}=24.1+0.182 \times \text{SYSBP}-0.064 \times \text{BALANCE}-0.0030 \times \text{VITAL}+0.111 \times \text{FAT}$

4) The high reliability was verified by analysis of residuals. And, the validity of this equation was verified by the results of preceding studies.

5) When MAP was added to these independent variables, at the 1st step MAP entered into the equation by the forward stepwise method ($F_{\text{ENTER}}=2.0$). SYSBP at the 2nd step, and BALANCE at the 3rd step were respectively entered. In this case, multiple correlation coefficient was 0.773. However, the calculated MAP was unequal to other variables as for the indicator of age changes because this variables were essentially designed to correlate to chronological age.

(Journal of Health Science, Kyushu University, 7:1~10, 1985)

緒 言

発育期にいわれる形態学的年齢・二次性徴年齢・歯芽年齢などはそれぞれの徴標の発育状態を基準に暦年齢にたちかえるものであり、年齢を予測する逆問題の一つの解とみなすことができる。青年期以降の加齢に関する研究はそれぞれの学問領域が固有な変化を用いて年齢基準値を設定するような場合が多く、人間という種々の機能を有し、かつまたそれらが複雑に関連し合っている系(システム)としての加齢が論じられることは少ないようである。その背景には「加齢変化の起こる時期」が各機能によって多様であり⁹⁾、「変化の勾配も小さく」⁶⁾、度数分布の偏移¹⁹⁾²⁰⁾など種々の特徴が考えられよう。

人類学・法医学的立場では歯芽の咬耗²⁴⁾²⁵⁾、循環器系や体力的変数⁶⁾²⁰⁾²⁹⁾、頭蓋や恥骨の縫合状態など⁷⁾¹³⁾を基準に重回帰分析やカテゴリカル回帰分析(林の数量化理論I類)といった多次元、多変量解析手法により老化度を測ろうとする試みもいくつか存在する。それらを概観すると特に重回帰分析の一次的終結である残差変動の分析が不十分であったり、標本の性別が不明であるような場合も少なくない。

体力学的な立場では本質的に運動動作を媒介したり、生理学的変数の導入を不可欠とする。田辺(1976)によると逆問題の解を求めるには「アприオリな知識を利用し、漸次真の構造に近づか」ねばならず、「制御しきれない誤差によってそれらの問題が発生していれば統計的モデルに帰着するのは自然である」²⁶⁾とされる。

本研究はこれらの問題所在を背景に青年期から初老期に至る婦人の暦年齢を体力学、内科学、形態学の計測値から予測するための回帰方程式(以下、単回帰方程式・重回帰方程式を含め、特に断らない限り回帰式と略す)の作成を意図したものである。特に変数選択には最近の統計学の知見を導入し、様々の視点からの考察を行おうとするものである。

研究方法

1. 対象

昭和56年度九州大学特定研究(「生活形態と健康度に関する総合的研究」)に被験者として任意に参加した健康な、大学女子教職員8名ならびに福岡市内の住宅地に居住する主婦40名総計48名を対象とした。被験者の暦年齢の平均は40.5歳、範囲は23歳から59歳までであった。

2. 独立変数

内科学的に収縮期血圧[コードSYSBP]、血中のクレアチニン[CREAT]、やアルブミン[ALBUM]の定量値を用いた。血中成分の分析法はクレアチニンをJaffe変法、アルブミンをBCG法により行った。

また形態学的変数として皮下脂肪厚(以下皮厚と略す)のうち上腕背部と肩甲骨下部とを栄研式皮脂厚計で測定し、その和[FAT]を用いた。

総合的な体力には下位構造として筋力や器管の強さ、柔軟性、運動協調性、持久性、速さなどがあり、さらにその下位構造が存在するという点が最近の多因子構造を仮定した研究¹⁰⁾でも首肯されている。また測定方法の客観性、基準値の明確性などの要件を踏まえた上で、研究の目的に照らさせ加齢による変化が明瞭である変数を採用すべきであろう。以上の論点より筋力では握力[GRIP]、背筋力[BACK]、器管の強さとして肺活量[VITAL]、運動協調性ではさらに下位構造を想定し平衡性として閉眼片足立ち[BALANCE]、柔軟性として立位体前屈[FLEX]を考え、一般的な方法²⁷⁾で測定した。

体力は仮説構成体としての概念にすぎないが、従来の研究に依拠すれば運動協調能にはその他協応性や敏捷性の側面も考えられるべきであり、また重要な持久性の側面、有酸素的体力については等閑視していることも見逃すことができない研究上の問題である。その意味でこれらのテストだけでは体力の全領域的測定が行われたとは、無論、いい難い。

3. 重回帰分析

暦年齢 (y_i) と 9 個の独立変数 (x_1, x_2, \dots, x_9) の測定値 ($N=48$) との関係線を線型モデル $\hat{y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_9 x_{9i}$ で定式化し、「推定値 \hat{y}_i と実測値 y_i との差の平方和」を最少にする $\beta_0, \beta_1, \beta_2, \dots, \beta_9$ を求めるため変数増加方式、変数増減方式によるステップワイズ重回帰分析を試みた⁴⁰⁾。

ここでデータ数について付言することにする。小林 (1982) は重回帰分析におけるデータ数について分散分析表にまとめた時の残差の自由度として「少なくとも 10 以上、できれば 50 以上もあることが望ましい」と述べている¹²⁾。本研究で対象とするのは 48 名であって、全ての独立変数を投入した時の回帰分析の自由度は 38 となり、望ましいとされる水準には達していないが、過少と判断することもできない。

データ数の多少以上に回帰分析自体の精度は残差分析によって保証されるべきものであり、これ以上の論議は必要としないかもしれない。

結果ならびに考察

1. 基本統計量

Table-1 は測定値の平均ほか基本統計量を示したものである。このうち *SKEW*, *KURT* は 4 次までの積率によって直接計算した統計量であり、それぞれ正規分布に比した歪度, 尖度を示す。特にこれに注意すると, *CREAT*, *BALANCE* の歪度は無視できない程度に大きい。

2. 相関係数

Table-2 には従属変数を含めた 10 変数間の相関係数を示す。独立変数間の相関係数に限定すると総計 36

個の相関係数のうちで有意となる ($P < .05$) のは 7 個に過ぎない。独立変数間の相関係数が高い場合には多重共線性 (multi-co-linearity)³¹⁾²²⁾ の問題がある。

チャタジーとプライス³⁾ (1977) は多重共線性の手掛りに①行列固有値に 0.01 未満のものが存在したり, ②固有値の逆数和が独立変数個数の 5 倍以上になる場合をあげている。

独立変数間相関行列 (9×9) の固有値を求めたところ, $\lambda_1 = 2.369$, $\lambda_2 = 1.699$, $\lambda_3 = 1.216$, $\lambda_4 = 0.956$, $\lambda_5 = 0.854$, $\lambda_6 = 0.638$, $\lambda_7 = 0.483$, $\lambda_8 = 0.457$, $\lambda_9 = 0.325$ となり, これらの逆数和は 12.964 となった。つまり上記①, ②のいずれの条件に照らしても多重共線性には問題がないことになる。なお, 多重共線性への対処について小林 (1982) は変数増減法の有効性を述べ¹²⁾, 佐和 (1979)²¹⁾ はリッジ回帰分析にそれを認めている。

3. 変数の投入と回帰式の評価

回帰式に投入される変数の順位は単に従属変数との関連のみでなく, 投入済み変数との関係で決定される。最初のステップ I では従属変数との単相関が最も高い変数が投入されるが, ステップ II 以降では次のような手順で変数を順次, 投入していった。すなわちステップごとに未投入変数が k 個あるとし, それらの個々を次のステップで投入したと仮定して回帰分析を行い, その時の F 値が最大となる変数を当該ステップで実際に投入する手続きを行い, これを繰返した。この時の F 値を *ENTER* 値とし, 次のように定義される⁴⁾。

$$FENTER = (RSSi' - RSSi + 1) / (RSSi + 1) / (N - P' - 1)$$

[ここで $i = 1, 2, \dots, m$ はステップ数, $N =$ 標本数,

Table 1. Condescriptive statistics

VARIABLE		MEAN	STD	MIN	MAX	SKEW	KURT
AGE	(yrs.)	40.5	8.1	23.0	59.0	0.237	-0.580
SYSBP	(mmHg)	121.0	17.4	90	162	0.533	-0.129
CREAT	(mg/dl)	0.804	0.107	0.60	1.10	0.835	0.566
ALBUM	(%)	64.2	2.96	56.9	69.1	-0.327	-0.731
GRIP	(kg)	28.9	4.74	17.0	38.0	-0.247	-0.473
BACK	(kg)	68.7	15.5	40.0	106.0	0.191	-0.482
BALANCE	(sec)	56.7	49.8	4	120	1.040	0.409
FLEX	(cm)	10.9	6.6	-1.8	24.1	-0.038	-0.943
VITAL	(cc)	2456	524.5	1200	3860	0.362	0.011
FAT	(mm)	37.9	12.3	15.5	62.5	0.068	-0.745

Table 2. Pearsonian correlation coefficient matrix(lower left).

	AGE	SYSBP	CREAT	ALBUM	GRIP	BACK	BALANCE	FLEX	VITAL	FAT
AGE	---									
SYSBP	559**	----								
CREAT	-005	032	----							
ALBUM	-250*	-397**	-234	---						
GRIP	-174	026	214	-126	----					
BACK	-299*	-162	044	082	452**	----				
BALANCE	-464**	-129	-115	-170	182	185	----			
FLEX	-067	-204	129	173	088	076	-027	---		
VITAL	-378**	-303*	035	036	449**	549**	173	156	---	
FAT	336**	397**	-087	-330*	-187	-250*	034	-067	-150	---

decimals omitted, ** $P < .01$, * $P < .05$

P' = 投入済み変数に該当変数を加えた個数, RSS = 当該ステップの回帰分散分析の残差平方和]

また一度投入された変数でも他の変数が追加投入されることにより, 回帰式の中で有意な貢献をしなくなり得る場合もある。そこで投入済み変数については $FREMOVE$ 値を求めた。

$$FREMOVE = (RSSi' - RSSi) / (RSSi) / (N - P')$$

ここで $RSSi'$ は「仮りに該当変数を除去すると仮定した時の回帰分散分析の残差平方和」であり, その他の記述は上記に従う。

ここでは $FENTER$ 値を手掛りに変数を次々増加していく変数増加方式と $FENTER$, $FREMOVE$ 両値を並行して用い, 投入・除去を繰返す増減方式の二つを行うことにした。

変数増加方式では次のような順序で変数が投入された。①SYSBP ($FENTER = 21.1$) → ②, BALANCE ($F = 13.3$) → ③VITAL ($F = 2.34$) → ④FAT ($F = 2.09$) → ⑤ALBUM ($F = 1.16$) → ⑥CREAT ($F = 0.58$) → ⑦FLEX ($F = 0.40$) → ⑧BACK ($F = 0.05$) → ⑨GRIP ($F = 0.01$)

得られた重回帰方程式は次式(1)のように記述され, 重相関係数 R は 0.738, 寄与率 54.5%, 推定値の標準誤差 SEE は 6.1 (歳) となる。

$$\hat{y} = 63.5 + 0.164 \times SYSBP - 0.0690 \times BALANCE - 0.0020 \times VITAL + 0.078 \times FAT - 7.19 \times ALBUM - 0.463 \times CREAT + 0.087 \times FLEX - 0.015 \times BACK - 0.0090 \times GRIP \dots (式1)$$

ところで P 個の独立変数があればそれらが構成する方程式個数は単回帰の場合の回帰式数 P から全変数が含まれる回帰式 (回帰式数 1) まで全てを考慮せねばならないので $2^P - 1$ 個にのぼる。本研究では

$2^9 - 1 = 511$ 個の方程式が考えられる。回帰式の評価はこれらの全てを逐一検討することで達成されようが, 合理的とはいえない。また実用面を考慮すれば単に独立変数が多いことが良好な方程式ともいえない。田辺 (1976) によると, 「(パラメータ数が) 比較的大きいモデルはデータへの見かけ上のあてはまりは良くなるが, データに含まれる誤差に過剰に反応してしまうため安定した構造情報をとりだすことができず, 「逆に自由度が小さすぎるモデルは現象の構造を十分反映できなくなり, 得られた情報に偏りを生ずる」ことになるとし, 「安定性と偏りを妥協させるモデルの評価基準が必要である」と述べている²⁰⁾。

この点で重相関係数 R は変数追加に従って上昇するので評価基準としては好ましいものではない。そのため, 自由度調整済み重相関係数 R^* , C_p 基準を算出し, また「モデルが定めた確率分希の“真”の確率分布への近似度の問題」として統計モデル選択を考えた赤池氏の情報量規準 AIC ¹²⁾¹¹⁾²⁴⁾²⁵⁾ を求めて回帰式の評価を行うことにした。

先述の変数増加方式による回帰分析では逐次変数を投入していった結果, 合計 9 個の回帰方程式が得られている。Table-3 は $FENTER$ 値によって選択・投入されていった変数名とそれらを含んだ回帰式の R^2 , R^{*2} , C_p , AIC を示している。また Fig-1 はこれらの変化を明示するためのものである。

AIC は「あてはまりの悪さ」の指標と考えられ, 小さいほどモデル (回帰式) 自体はすぐれているといえる¹²⁾¹¹⁾²⁴⁾²⁵⁾。 C_p でも同様なことが考えられる。

この結果, AIC では変数 4 個を含んだステップ IV, C_p ではステップ III の回帰式を各々, 最良と評価して

いる。 R^{*2} では R と同じくこの増大があたりはまりの良さを示すが、ステップVで R^{*2} は最大となり、 R^2 は最終ステップIXまで増大を続ける。

AICはその後、ステップVからは上昇を開始し、また C_p ではステップIV以降で上昇を始め、 R^{*2} はステップIX以降から低下を示している。つまりAICでは4変数 [SYSBP→BALANCE→VITAL→FAT] からの回帰式を最良と評価し、 C_p はこれから [FAT]

を除いたものを、また R^{*2} はAICの場合の4変数に [ALBUM] を加えた5変数からなる回帰式をそれぞれ最良と評価することになる。

母数節約的²⁰⁾にモデル [回帰式] 構築を行うとすれば C_p による場合が良く、分担した内科学、体力学、形態学からいえば R^{*2} に依拠するのが好ましいと考えられる。AICは両者の折衷点にあり、適切かつ合理的な変数選択が行われていると評価できるであろう。

Table 3. Entered variables and several criteria statistics.

STEP	VARIABLE	R^2	R^{*2}	AIC	C_p
I	+SYSBP	0.313	0.299	183.3	13.270
II	+BALANCE	0.470	0.446	173.1	2.226
III	+VITAL	0.497	0.462	172.6	1.994
IV	+FAT	0.521	0.476	172.3	2.050
V	+ALBUM	0.533	0.478	173.0	2.976
VI	+CREAT	0.539	0.472	174.6	4.435
VII	+FLEX	0.544	0.465	175.9	6.052
VIII	+BACK	0.545	0.452	177.8	8.001
IX	+GRIP	0.545	0.437	179.8	10.000

R^2 , multiple correlation coefficient

R^{*2} , multiple correlation coefficient adjusted for d.f.

AIC, Akaike's an information criterion

C_p , Mallows' criterion

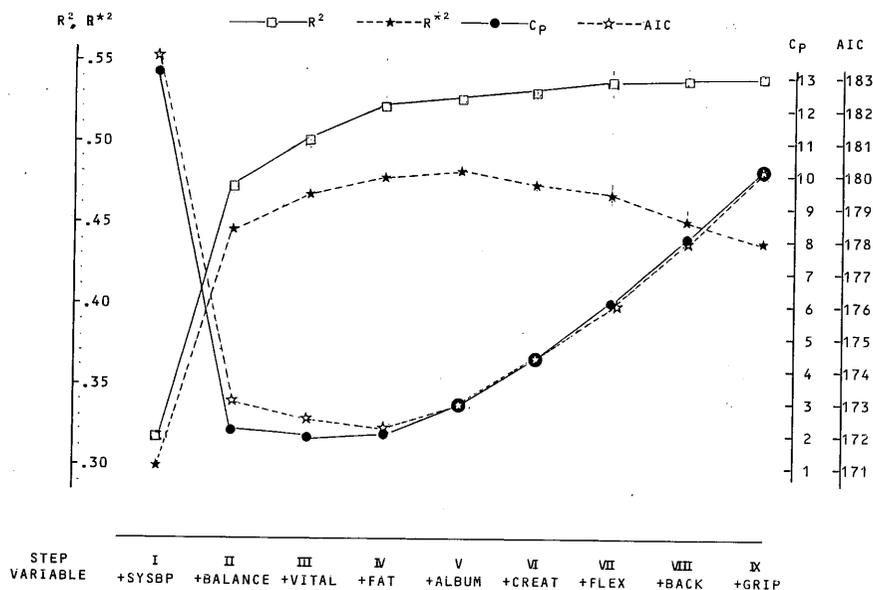


Fig. 1. The changes of criteria statistics and entered variables on each step.

すなわち変数増加方式では SYSBP, BALANCE, VITAL, FAT と変数を増していき、ここまでで変数投入を打切っても、なお被験者数と変数個数の関連で十分に妥当性の高い、年齢予測の回帰式が得られるといえよう。

次に $FENTER \geq FREMOVE$ 値を併行させながら算出し、変数の投入と除去を繰返した変数増減方式の回帰分析を行った。 F 基準値 ($FENTER, REMOVE$) の設定については研究者間で違いがあり、概ね 1.0 ~ 3.7 程度が考えられている³⁵⁾¹⁷⁾²⁰⁾。本研究では F 基準値として 2.0 を設定し変数の投入、除去を行うことにした。

この結果、①SYSBP→②BALANCE→③VITAL→④FAT の順で変数投入が行われ、この後には追加投入される変数すなわち F 基準値 ($F \geq 2.0$) を満足する変数は見当らず、また①から④までのステップで投入された変数は中途の過程で $FREMOVE$ 値が 2.0 を下回ることもなかった。すなわち中途での変数の除去は行われなかった。参考値としてあげると、ステップ V を仮に考えた時、投入される予定の変数は ALBUM である。この $FENTER$ 値は 1.16 であり、ステップ IV で $FREMOVE$ 値が最小となるのは FAT ($F = 2.09$) であって両者の $FENTER, FREMOVE$ 値間には大きな差が認められる。

井上などの指摘では AIC と Cp を予測の適中率からみた時、両者間に有意差は認められない¹⁰⁾といわれる。本研究では両基準間で選択された変数個数に差異が認められた。赤池 (1981)²⁾ は「奥野ほか (1976)¹⁷⁾ が示す変数増減法はつまるところ AIC 最小化に対応したものであることを述べ、その具体的考察は杉山ほか (1979)²⁴⁾ にみることができ。本研究の結果として重回帰分析における F 基準値が 2.0 であれば AIC 最小化も果たし得ることが具体的となった。すなわち F 基準値 2.0 とする変数増減法で構築された回帰式は AIC 最小化 ($MAICE$) をも果たすことになるといえよう。

得られた重回帰式は次のように表わされ、 $R = 0.721$, $SEE = 5.9$ 歳となり、全変数を含んだ式 (1) と比べても信頼性は劣るとはいえない。

$$\hat{y} = 24.1 + 0.182 \times \text{SYSBP} - 0.064 \times \text{BALANCE} - 0.0030 \times \text{VITAL} + 0.111 \times \text{FAT} \dots \text{(式 2)}$$

また回帰分析の結果は $F = 11.67$ ($P < .01$) となり、相応の信頼性を確認できた。

4. 独立変数の検討

回帰式 (2) の妥当性を検討するため各変数の標準化偏回帰係数 (以下、標準化係数) を手掛りに以下の

ように考察した。収縮期血圧 [SYSBP] は標準化係数 0.392 であるが、加齢とともに SYSBP は上昇するという点は先進工業国住民に関して一致している¹⁴⁾¹⁸⁾²⁰⁾²⁷⁾。体力変数のうち BALANCE, VITAL の標準化係数は、それぞれ -0.391, -0.165 であり、平衡機能は年齢とともに低下していくと考えられている。

肺活量はその減衰開始年齢には多少の差異が報告間にみられるものの、年齢との負の相関関係はそれらの研究でも首肯されている¹⁸⁾³⁰⁾。

形態学的肥満の指標に皮脂厚値 FAT を用いたが、標準化係数 0.170 となっている。特に女子では皮脂厚値は加齢とともに漸増し、60 歳前後までその傾向が続くとされている²⁰⁾²²⁾。

以上のように回帰式 (2) に含まれた変数はその一般的な知見と標準化係数、特に年齢との相関関係を照合しても矛盾点はないと考えられる。また体力変数、収縮期血圧、形態学的肥満に関する変数で構成されている点は分担した研究領域からしても興味あるものとして考えることができよう。

5. 残差分析

得られた回帰式の欠陥、換言すれば信頼性の確認は回帰分散分析のほか残差分析を必要とする。暦年齢の実測値 y_i と式 (2) によって得られた予測年齢 \hat{y}_i との差すなわち残差 e_i およびその標準偏差 Se などを利用して残差分析を行った。残差の統計量は Table-4 のようになる。残差の分布の正規分布に比して無視してもよい程度の歪みしかもたない。残差がその標準偏差 (Se) に対し $1Se$ を越えた場合は 3 名であった。

Fig-2 は標準化残差 e_{is} ($= (y_i - \hat{y}_i) / Se$) と標準化予測値とで構成された座標として個人をプロットしたものであり、ゼロの周りに不規則に分布し、特定の変動は認められない³⁾ようである。

以上より最良と評価した式 (2) には特に注意すべき欠陥はないものと考えられる。

6. 持久性指標の導入について

この種の研究、特に対象者が成人である点を重視すると持久性 (Aerobic power) の指標を導入することの意義や有効性については論を待たない。

本研究でも Margaria et al (1965)¹⁵⁾ の提唱した手続きで間接的な Maximal Aerobic Power (MAP) を算出し、基本統計量の算出から、重回帰分析に至る手続きを試みた。以下、明らかになった点、問題点を列挙すると次のようになる。

(1) MAP の歪度は 1.502, 尖度は 2.905 となり、この値は正規分布に比し、必ずしも無視できる程の値で

Table 4. Condescriptive statistics of residuals.

S.D.	25%ile	MED.	75%ile	SKEW
5.52	-4.0	-1.2	3.2	0.654 ($t=1.05, P<.05$)

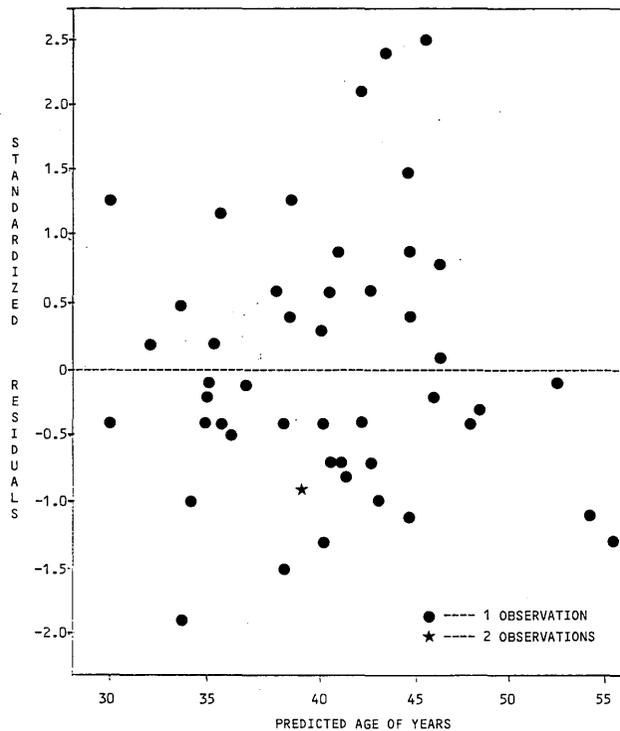


Fig. 2. Plottings of standardized residuals and predicted age.

はない。

(2) MAPの算出が心拍数と年齢との間に、一次線型関係を同定して行われているので、ステップワイズ方式では第1変数として投入される。

(3) F 基準値2.0と設定すると、MAP→BALANCE→SYSBPの順に回帰式に投入され、ここまでで変数の選択は打ち切られた。また、ここまでのステップでAIC, Cpともに最小となった。この方程式は $R=0.773$, $SEE=5.25$ 歳となり、次のように表現できる。

$$\hat{y} = 52.65 + 0.140 \times \text{SYSBP} - 0.801 \times \text{MAP} - 0.064 \times \text{BALANCE}$$

(4) MAPを除いた変数で構築され、最良と評価された式(2)にMAPを追加投入したところ、

$$\hat{y} = 57.9 + 0.129 \times \text{SYSBP} - 0.014 \times \text{FAT} - 0.001 \times \text{VITAL} - 0.062 \times \text{BALANCE} - 0.866 \times \text{MAP}$$

と表現され、 $R=0.754$, $SEE=5.6$ (歳)となった。(5) この回帰式に含まれる変数のうちFATは負の回帰係数を示し、この点はFATに関する一般的な知見とは符号しない。

これらの点を鑑みると、間接的な算出値特に年齢との一次的な関係を前提としたMAPをそのまま用いることは、この種の年齢研究では必ずしも適さず、むしろ古川ほかの一連の研究⁹⁾²⁸⁾に示されるように心拍数そのものを利用して持久性の指標化を企図することが必要と考えられる。

結論ならびに討論

23歳から59歳までの健康な女子48名(平均年齢40.5歳)について収縮期血圧[SYSBP], 血中クレアチニン[CREAT], 血圧アルブミン[ALBUM]の各定量値, 皮脂厚[FAT], 握力[GRIP], 背筋力[BACK], 立位体前屈[FLEX], 閉眼片足立ち[BALANCE], 肺活量[VITAL]から暦年齢を予測する重回帰方程式の作成を試みた。

結果は以下の様に要約できる。

1) CREAT, BALANCEの歪度は無視できない程度に大きく, 独立変数間相関行列を検討した時, 多重共線性は示唆できなかった。

2) 変数増加方式で9個の独立変数を逐次投入する重回帰分析では①SYSBP→②BALANCE→③VITAL→④FAT→⑤ALBUM→⑥CREAT→⑦FLEX→⑧BACK→⑨GRIPの順に投入されていき, この場合, $R = 0.738$, $SEE = 6.1$ 歳となった。

3) この9ステップを独立と考えた9個の回帰式を自由度調整済み重相関係数 R^* , 赤池の情報量基準AIC, C_p 基準で評価したところ C_p はSYSBP, BALANCE, VITALの3変数を含んだ回帰式, AICは第4変数としてFATを加えた回帰式, R^* は第5変数としてさらにALBUMを加えた回帰式をそれぞれ最良と評価した。

4) $FENTER = FREMOVE$ の基準値を2.0とする変数増減方式でもAICによって最良と評価された4変数による回帰式が採択され, $R = 0.721$, $SEE = 5.9$ 歳となり, 次のように記述される。

$$\hat{y} = 24.1 + 0.182 \times \text{SYSBP} - 0.064 \times \text{BALANCE} - 0.0030 \times \text{VITAL} + 0.111 \times \text{FAT} \dots \dots \text{(式2)}$$

5) AICで最良と評価されたこの回帰式に含まれる4変数の意義や符号の検討などの考察によると, いずれも妥当と考えられるものであった。またこの式(2)の残差分析でも回帰式の欠陥は特に認められず, 残差の標準偏差は5.52となり, 正規性検討などでも問題は考えられないようである。

〈附記〉

本研究は昭和56年度, 九州大学特定研究「生活形態と健康度に関する総合的研究」として行われたものの一環である。この研究には健康科学センターのスタッフ全員が係り, データを集積した。執筆は著者があたり, データ解析も, 九州大学大型計算機センターにお

いて著者自身が担当した。

いうまでもなく, 本研究のデータは健康科学センター全スタッフの共有のものであり, 執筆の責は著者自身にある。その意義からして直接に検査・測定にあたられた全スタッフに深甚なる謝意を表せねばならない。

また被験者として参加された九州大学教養部の教職員の皆さん, 福岡市東区八田青葉台団地の住民の皆さんに衷心より御礼を申しあげたい。

本研究の一部には九州大学大型計算機センターの英文清書システム(ROFF)を利用した。

文 献

- 1) 赤池弘次: 情報規準 AIC とは何か——その意味と将来への展望——, 数理科学, 14(3): 5-11, 1976.
- 2) 赤池弘次: モデルによってデータを測る, 数理科学, 19(3): 7-10, 1981.
- 3) チャタジー・プライス: 回帰分析の実際(佐和隆光・加納悟訳), 新曜社, 1981, Pp.249.
- 4) Dixon, W. J. and M. E. Brown (Eds.): BMD—Biomedical computer programs P-series, Univ Calif. Press: Berkley, 1979, 399-417.
- 5) ドレーパー・スミス, 応用回帰分析(中村慶一訳): 森北出版, 1968, 163-216.
- 6) 古川俊之・刈田全世・吉川博通・稲田紘・森脇健・阿部裕・宮崎学・梶谷文彦・宮脇一男: 多変量理論による老年者の生物学的年齢推定, 高齢医学, 6: 308-317, 1968.
- 7) Hanihara, K. and T. Suzuki: Estimation of age from the pubic symphysis by means of multiple regression analysis, Am. J. Phys. Anthropol., 48: 233-240, 1978.
- 8) 芳賀敏郎・橋本茂司: 回帰分析と主成分分析, 日科技連, 1980, Pp. 228.
- 9) 菱沼從尹: 年齢の断面, 発達1(3): 10-18, 1980.
- 10) 井上隆勝: 線型重回帰モデルにおける一つのモデル選択規準, 応用統計学11(2): 63-80, 1982.
- 11) ケンドール: 多変量解析(奥野忠一・大橋靖雄訳): 培風館, 1981, Pp. 225.
- 12) 小林龍一: 相関・回帰分析入門(新訂版), 日科技連, 1982, Pp. 259.
- 13) Koizumi, K: The estimation of age from the cranial sutures by means of multivariate

- analysis methods, *J. Anthropol. Soc. Nippon*, **90** (2): 109-118, 1982.
- 14) 葛代隆・田淵義勝・矢野敦雄・小沢秀樹・池田正男：食塩——食塩と高血圧について，*内科*，**48**(5): 749-752, 1981.
- 15) Margaria, R., P. Agehemo, and E. Revelli: Indirect determination of maximal O₂ consumption in man, *J. appl. Physiol* **20**(5): 1070-1073, 1965.
- 16) 松浦義行：運動能力の因子構造，不昧堂，1969, 89-101.
- 17) 奥野忠一・久米均・芳賀敏郎・吉澤正：多変量解析法，日科技連，1978, 7-76.
- 18) 小野三嗣：健康と体力の科学，不昧堂，1971, 189-204.
- 19) 大島正光：ヒト——その未知へのアプローチ，同文書院，1982, 226-239.
- 20) 小沢利男：老年者の循環機能，*日本臨床*，**39**(3): 510-514, 1981.
- 21) 佐和隆光：回帰分析，朝倉書店，1979, Pp. 187.
- 22) Shephard, R. J. : *Physical activity and aging*, Croom Helm, London, 1978, 52-145.
- 23) 白石透・毛利昌史・本田憲業：老年者の呼吸機能——加齢による変化と老年者における評価——，*日本臨床* **39**(3): 510-514, 1981.
- 24) 杉山高一・尾崎公・牛沢賢治・清水政利：歯の咬耗度による年齢推定に関する重回帰分析，*応用統計学*，**5**(3): 123-128, 1979.
- 25) 竹井哲司：歯の咬耗による年齢の推定，*日法医誌* **24**(1): 4-17, 1970.
- 26) 田辺國士：不適切問題への統計的アプローチ，*数理科学*，**14**(3): 60-64, 1976.
- 27) 東京都立大学身体適性学研究室編：日本人の体力標準値Ⅲ版，不昧堂，1980, Pp. 458.
- 28) 柳井晴夫・岩坪秀一：複雑さに挑む科学，講談社，1976, 156-176.
- 29) 吉川博通：多変量理論による生物学的年齢の推定，*阪大医誌* **22**(1, 2): 113-136, 1976.
- 30) 吉川政己：老化度・健康度をどう測定するか，*内科*，**48**(5): 729-735, 1981.