女子学生の1000m走の回帰分析

青山昌二

Regression analysis of College female student's 1000 meter run

by

Shoji Aoyama

The purpose of this study is based upon computing the mnltiple regression equation, which predicts the 215 College female student's 1000 meter run from their 29 physique and physical fitness measurement items. The specific point of this equation means that selection order of the predictor is defined according to going up more and more the order of partial correlation coefficient.

The results were as follows;

(1) Eleven items were selected for the predictor, and the other items as predictor were not significant for criterion of high order partial correlation coefficient.

(2) Multiple correlation coefficient was R = .631.

These eleven items and the order were 1) running broad jump, 2) step test, 3) pull-ups, 4) side step test, 5) span, 6) maximum girth of forarm, 7) standing trunk flexion, 8) vital capacity, 9) body height, 10) hand ball throw and 11) back strength.

1 研究目的・方法

この研究は，女子学生（体育専攻，215人に対して実施した29項目の体格・体力・運動能力測定結果の相関係数マトリクス（昭和45年測定）[1]から，1000m走を基準変数，その他の測定項目を予測変数とする重回帰方程式を作成し，それによって，1000m走に対する他の28項目による推定能力を吟味すること，およびそのさい予測変数の採択において，どのような順序で選択していったら推定能力の上昇率からみて最も効率的なものとなりうるかを，偏相関によって消去次元を高めていくことによって吟味し明らかにすること，さらには，こうした重回帰方程式がいわゆる測定一評価という観点で捉えて，どのような意味をもつものであるかということについても若干の考察を試みること，をねらいとするものである。

2 結果および考察

(1) 予測変数の選択順序

もちろん，1000m走以外の28測定項目すべてを予測変数とする回帰方程式を作成して，重相関係数を調べることは可能であるが，ここで問題にしようとしていることは，できるだけ少ない個数の予測変数で効果を最大にする，つまり重相関係数を高めるにはどのように予測変数を選択していけばよいか，ということである。このための方法として，1000m走との偏相関係数の次元を漸増していく，という方法を採用することとする。

表1はその結果である。まず，1番目の予測変数として，1000m走との相関係数の最も大きい(絶対値でみて)，r＝−0.400 の走り幅とびを選択する。標準得点で表わせば

$$\hat{Y} = -0.400\,X$$

表1　消去変数の数を順次増していった場合の各測定項目と1000m走との偏相関係数

		偏相関係数（消去変数の個数）										
	1000m走との相関係数	1	2	3	4	5	6	7	8	9	10	11
1 身長	−.070	−.029	−.039	−.094	−.080	.086	.072	.077	$\boxed{.136}$			
2 体重	.096	.034	.033	.002	.008	.106	−.022	−.017	.007	−.048	−.040	−.039
3 胸囲	.033	−.010	.002	.004	.006	.061	−.055	−.037	−.029	−.027	−.025	−.012
4 座高	.035	.041	.059	.014	−.006	.079	.045	.033	.082	−.007	−.003	.008
5 頸囲	.058	.035	.040	.047	.051	.106	.030	.032	.037	.019	.020	.024
6 腰部最小位	.103	.027	.025	−.001	−.008	.041	−.065	−.034	−.035	−.030	−.015	−.020
7 上腕伸囲	.110	.022	.026	.045	.054	.078	−.064	−.065	−.083	−.083	−.083	−.083
8 前腕最大囲	.088	.038	.078	.090	.108	$\boxed{.150}$						
9 大腿最大囲	.137	.035	.042	.019	.006	.036	−.111	−.108	−.093	−.107	−.111	$\boxed{.085}$
10 下腿最大囲	.126	.071	.070	.069	.077	.121	.022	.007	.016	−.006	.002	−.002
11 指極	−.128	−.096	−.128	−.170	$\boxed{-.156}$							
12 肩幅	−.137	−.097	−.104	−.127	−.115	−.019	−.067	−.067	−.046	−.054	−.062	−.055
13 骨盤幅	.061	.035	.016	−.016	−.029	.025	−.020	−.021	−.013	−.044	−.044	−.046
14 腹部皮厚	.223	.136	.138	.106	.122	.116	.068	.085	.076	.074	.065	.063
15 肺活量	−.178	−.151	−.168	−.175	−.152	−.101	−.132	$\boxed{-.154}$				
16 反復横とび	−.260	−.198	−.200	$\boxed{-.182}$								
17 垂直とび	−.210	−.050	−.067	−.034	−.021	.012	.019	−.005	−.007	−.019	−.022	−.027
18 背筋力	−.059	.017	.052	.073	.079	.123	.092	.071	.093	.105	$\boxed{.112}$	
19 握力(右)	−.046	.000	−.009	.027	.032	.087	.026	.014	.016	.009	.007	−.028
20 握力(左)	−.074	−.039	−.032	−.009	−.015	.049	−.014	−.030	−.015	−.004	−.002	−.040
22 踏台昇降	−.313	$\boxed{-.334}$										
23 伏臥上体そらし	.089	.124	−.092	.117	.130	.124	.129	.066	.079	.041	.043	.044
23 立位体前屈	.152	.099	−.079	.134	.106	.119	$\boxed{.136}$					
24 50m走	.293	.068	.088	.047	.046	.032	.008	.011	.010	.012	.013	.025
25 走り幅とび	$\boxed{-.400}$											
26 ハンドボール投げ	−.270	−.164	−.155	−.161	−.127	−.102	−.131	−.126	−.128	$\boxed{-.118}$		
27 斜めけん垂	−.300	−.242	$\boxed{-.210}$									
28 片足立片足爪先立ち	−.009	.038	.035	−.045	.085	.082	.079	.058	.069	.070	.103	.082

である。

つぎに，こんどは走り幅とびの変数を消去して1000m走と他の変数との偏相関係数を各項目についてそれぞれ計算し，このうちで最も大きなもの（絶対値で）を，第2番目の予測変数として選択する。すなわち，第2番目の変数の選択は，1000m走との関係を第1番目の変数によってはカバーできない部分のうち，どの変数を選択すればそのカバーする量が最も大きくなるか，ということをみて選択しているわけである。第2変数は r＝−0.334 を示す踏台昇降である。

同様な考え方から，第3変数の選択では，第1・第2変数，すなわち走り幅とびと踏台昇降を消去したうえで，1000m走との偏相関係数（2次の偏相関係数）を残りの26の種目について求め，この中で最も大きい斜めけん垂（r＝−0.210）が選ばれる。

第4変数は第3変数までを消去した上でみた偏

相関係数の最も大きい反復横とび，第５変数は第４変数までを消去した上でみた偏相関係数の最も大きい指極，というように次々と選ばれる。すなわち，それまで選択された予測変数でカバーできなかった部分について，残りの変数のうちから最大カバー能力をもつ変数が選択されるという方法をとっていき，そのカバー能力が0.1を下まわったところで打ち切ることにした。こうして，選択された予測変数の順位とそれに伴う各次元の偏相関係数を表１から再度整理してみると表２のようになる。

これをみると，1000m走との相関係数は，踏台昇降（ｒ＝－0.313）よりも走り幅とびの方が高く，第１予測変数の位置にあること，身長と1000m走との相関は，単相関では ｒ＝－0.070と有意でないにしても数値の上ではマイナスをもっていたのに，それまでの８個の予測変数を消去してみると，プラスの値をとっている，つまりそれまでの８個の予測変数が〈固定されている〉とすれば身長の大きい方が不利である（といってもこの偏相関係数値ではその程度がわずかであるが）ということ，身長にみるこの傾向は背筋力と1000m走との間にもみられること，などが目につく。

⑵　重回帰方程式の導出

11個の予測変数とその順位が決定されたので，もう一度相関係数マトリクスからとり出して，この順位にしたがって予測変数の個数を２個から順次11個まで，10個の重回帰方程式を導出した。

$$\hat{Y} = \beta_1 X_1 + \beta_2 X_2$$

$$\hat{Y} = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\hat{Y} = \beta_1 X_1 + \beta_2 X_2 + \cdots\cdots + \beta_{10} X_{10}$$

表３に各式の標準偏回帰係数[2]（β）および重相関係数（R）の値を示す。

まず，重相関係数の値の上昇についてみると，1000m走と走り幅とびの単相関係数 ｒ＝0.4（絶対値で）から，走り幅とびと踏台昇降の同時２変数となると相関係数（R＝0.504）はほぼ0.5となり，予測変数の増加とともに相関係数（R）もわずかずつ上昇していって，予測変数の個数が７個になると相関係数（R＝0.595）がほぼ0.6となる。それからあとは重相関係数の上昇は一そう小さくなり，予測変数がいちばん多い11個のばあいでもR＝0.631である。これ以上予測変数の数を多くしたとしても，残りの種目と1000m走の偏相関係数からみて，重相関係数の上昇はほとんど望めない。R＝0.631を評価すると，それほど高い値であるとはいえない。しかし予測変数の数を11個まで多くしたことの効果はある程度認められる。

表２　1000m走の予測変数の選択順位と偏相関係数およびその次元

予測変数の選択順位	予測変数の種目名	1000m走との偏相関係数	偏相関係数の次元
1	走り幅とび	(－.400)	－
2	踏台昇降	－.334	1
3	斜めけん垂	－.210	2
4	反復横とび	－.182	3
5	指極	－.156	4
6	前腕最大囲	－.150	5
7	立位体前屈	－.136	6
8	肺活量	－.154	7
9	身長	.136	8
10	ハンドボール投げ	－.118	9
11	背筋力	.112	10

つぎに，各重回帰方程式の標準偏回帰係数（β）をみると，各標準偏回帰係数はその式における他の予測変数を〈固定〉して考えたときのその変数の基準変数1000m走に対する寄与率を示しているわけである[3]から，これによって予測変数どうしの相互関係によって決められてくる，基準偏数との関係の深さが〈相対的に〉(relative) わかる。加えて，各式の重相関係数値の２乗（回帰からの分散の，もとの分散に対する比）を100としたときのそれへの各予測変数のかかわりの程度（net contribution）をも（　）内に示した。

たとえば，予測変数２個の場合の重回帰方程式をみると，走り幅とびと踏台昇降の値をそれぞれ標準得点（$z = (X - \overline{X})/\sigma$）に変換して予測変数とする場合の，お互いに相手を〈固定〉して考えた上での係数が－0.394と－0.304で〈相対的に〉走り幅とびの方が踏台昇降より大きいということが示される。したがって，予測変数の数が多くなって，〈固定〉して考える変数の数が多くなると相対的な大きさが変わってくる。他の予測変数ではとって代わることのできない，いわば自分独自

表３　1000ｍ走を推定するための標準偏回帰係数と重相関係数

予測変数の数	重相関係数	走り幅とび	踏台昇降	斜めけん垂	反復とび	指極	前腕最大囲	立位体前屈	肺活量	身長	ハンドボール投げ	背筋力
2	.504	−.394 (62.2)	−.306 (37.8)									
3	.535	−.356 (49.7)	−.281 (30.7)	−.187 (19.6)								
4	.557	−.327 (42.2)	−.279. (28.1)	−.169 (16.4)	−.158 (13.3)							
5	.572	−.311 (38.1)	−.285 (27.3)	−.194 (17.8)	−.145 (11.6)	−.133 (5.2)						
6	.585	−.287 (33.5)	−.299 (27.4)	−.205 (18.0)	−.153 (11.7)	−.164 (6.1)	−.129 (3.3)					
7	.595	−.263 (29.8)	−.289 (25.5)	−.237 (20.1)	−.133 (9.8)	−.175 (6.3)	−.140 (3.5)	−.118 (5.0)				
8	.608	−.251 (27.2)	−.290 (24.6)	−.238 (19.3)	−.116 (8.1)	−.128 (4.4)	−.167 (4.0)	−.136 (5.6)	−.140 (6.7)			
9	.617	−.259 (25.2)	−.300 (22.9)	−.217 (15.9)	−.110 (7.0)	−.281 (8.8)	.164 (3.5)	.144 (5.3)	−.182 (7.9)	.209 (3.6)		
10	.624	−.229 (21.9)	−.295 (22.2)	−.219 (15.8)	−.092 (5.8)	−.256 (7.9)	.180 (3.8)	.139 (5.1)	−.179 (7.7)	.193 (3.3)	−.104 (6.7)	
11	.631	−.247 (22.6)	−.304 (21.7)	−.224 (15.4)	−.091 (5.4)	−.278 (8.1)	.159 (3.2)	.124 (4.3)	−.195 (7.9)	.206 (3.3)	−.109 (6.7)	.100 (1.3)

注）（　）内の数字はnet contribution を示す。

のもの（基準変数に対して）を多くもっている変数ほど，標準偏回帰係数は〈相対的〉に大きくなる。予測変数11個の場合をみると，踏台昇降が最も大きく，次が指極，その次が走り幅とびという順序になっている。

この式に，各人の実際に測定した値を標準得点であてはめ，その値（$\hat{Y}$）とその者の実際の1000ｍ走の値（標準得点）と比較して，どの面（種目）が劣っているか，Ｙをある一定量だけ高めるためには，どの面（どの種目）をどれだけあげていくとそれが可能となるかということの理論的な面をおさえることができる。実際には，統計的な割り出しでありつまり確率モデルであるので，個人レベルのＡという個人，Ｂという個人にそのまま通用するということはいえないにしても，目安としては用いることができる。具体的な指導の場において各人のＹを算出させ，その結果を各人に評価させたりして，各人の具体的な努力目標を立てさせることから，有効な指導のために役立たせることが可能となる。[4] 集団レベルでの指導においてはさらに有効なものとして評価することができよう。

注

1）青山昌二：「コンピュータによる計算」，和泉貞男著「体育統計」に所収，道和書院，1970，160～161頁。

2）芝祐順：「行動科学における相関分析法」，東京大学出版会，1967，17頁以下。

3）Willam W. Cooly & Paul R. Lohnes: Multivariate Procedures for the Behavioral Sciences, John Wily & Sons, 1962, p.50

4）青山昌二：「大学生の体格・体力の統計的分析」，「体育学紀要」第８号，東大教養体育研究室，1974，69—73頁。

青山昌二：「体育測定の指導と処理・その活用」，「体育の科学」第26巻第３号，1976，178—179頁。