

提出課題の相互評価に関する分析について

Analysis of peer evaluation of submitted assignments

小 西 敏 雄

Toshio KONISHI

(心理子ども学科子ども専攻)

要 約

ある演習科目で提出課題を相互評価させている。この結果をクラスター分析して分かったことを説明する。

キーワード：クラスター分析、分散分析、多重比較

[Abstract]

In a certain seminar subject, submitted assignments are evaluated by each other. I will explain what I found out by cluster analysis of this result.

Keywords: cluster analysis, analysis of variance, multiple comparison

1. 課題提出された作品の相互評価について

数年前から子ども専攻専門科目『幼児教育の方法』を影浦紀子先生と一緒に担当している。受講している約 50 人の学生を対象にして、私は最後の 5 回でビデオクリップ作りを教えている。

内容はパワーポイントのスライド自動切り替え機能を使って切り替え時間を設定し、動くスライドを作ることである。画像データの切り取り方から始めて、音楽データの CD からの取り方、パワーポイントの操作方法からスライドの作り方までを 4 回で教え、同時に作品を作っていく。最後の 1 回の前に作品を提出させ教室内サーバの共有領域に配置する。最終回の授業で受講学生に作品を見させて、自分以外の 49 人の評価をさせる。それを集めると、評価した側と評価された側のデータが得られる。

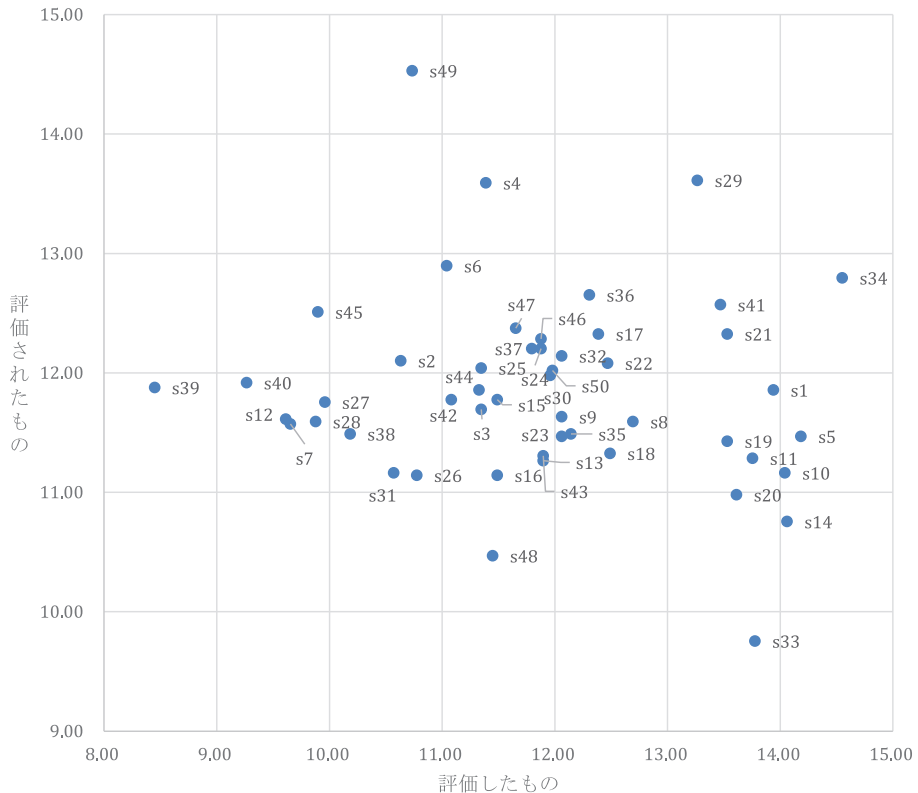
提出課題の相互評価に関する分析について

No.	評価した平均 mean1	評価したSD sd1	評価された平均 mean2	評価されたSD sd2
s1	13.94	0.98	11.86	1.75
s2	10.63	1.73	12.10	2.08
s3	11.35	0.85	11.69	1.80
s4	11.39	0.97	13.59	1.63
s5	14.18	1.06	11.47	1.86
s6	11.04	1.23	12.90	1.71
s7	9.65	1.24	11.57	1.84

例えば s1 の学生については、他の 49 人を評価した成績の平均が 13.94 で標準偏差が 0.98、他の 49 人から評価された成績の平均が 11.86 で標準偏差が 1.75 だということである。成績は 15 点満点である。すると、s1 の学生は、自分の評価は平均が 13.94 で標準偏差が 0.98 なので、割と高めに同じような点数をつけていることが分かる。ところが、評価された側は平均が 11.86 で標準偏差が 1.75 なので、評価はあまり高くなく、しかも自分がした評価よりもばらついていることがわかる。すなわち、評価した点数と評価された点数はあまり関連が無いことが分かる。

これを使って次のような散布図が得られた。

評価した×評価された平均の分布



散布図の中央部上位にある s49 の学生は他者に対する評価はすごく厳しいが、他者からは高く評価されていることが分かる。このような学生の存在が、相互評価には欠かせないと思われる。

次表は今回の評価する側・評価される側の平均と標準偏差で相関係数を計算したものである。

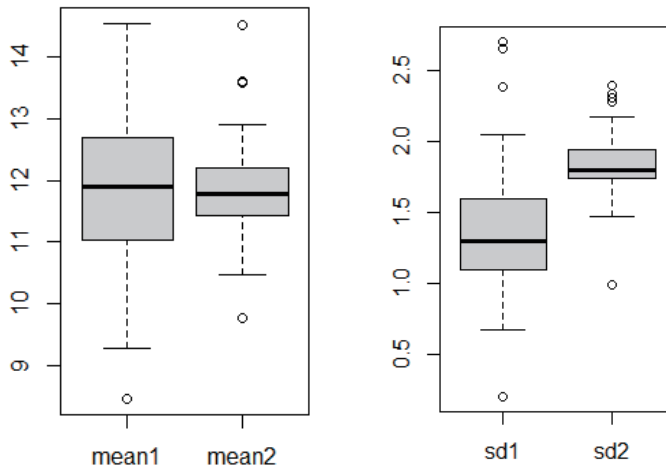
	mean1	sd1	mean2	sd2
mean1	1.000	-0.332	-0.122	0.023
sd1	-0.332	1.000	-0.055	0.058
mean2	-0.122	-0.055	1.000	-0.654
sd2	0.023	0.058	-0.654	1.000

mean1 と mean2 には当然のことながら、相関が無い。相関があった場合には何らかの事前相談や密談があったと予測される。mean2 と sd2 の相関係数は -0.654 で、負の相関が見られる。これは、成績の上限が 15 点なので、点数が上がるほど標準偏差が低くなることを意味するので、妥当な結果であると思われる。

今回のデータの基本統計量を計算した。

mean1		sd1		mean2		sd2	
Min.	: 8.449	Min.	: 0.1998	Min.	: 9.755	Min.	: 0.9918
1st Qu.	: 11.051	1st Qu.	: 1.0940	1st Qu.	: 11.439	1st Qu.	: 1.7394
Median	: 11.888	Median	: 1.3005	Median	: 11.776	Median	: 1.7967
Mean	: 11.858	Mean	: 1.3552	Mean	: 11.857	Mean	: 1.8359
3rd Qu.	: 12.643	3rd Qu.	: 1.5819	3rd Qu.	: 12.204	3rd Qu.	: 1.9352
Max.	: 14.551	Max.	: 2.7034	Max.	: 14.531	Max.	: 2.3905

今回のデータの箱ひげ図を作成した。



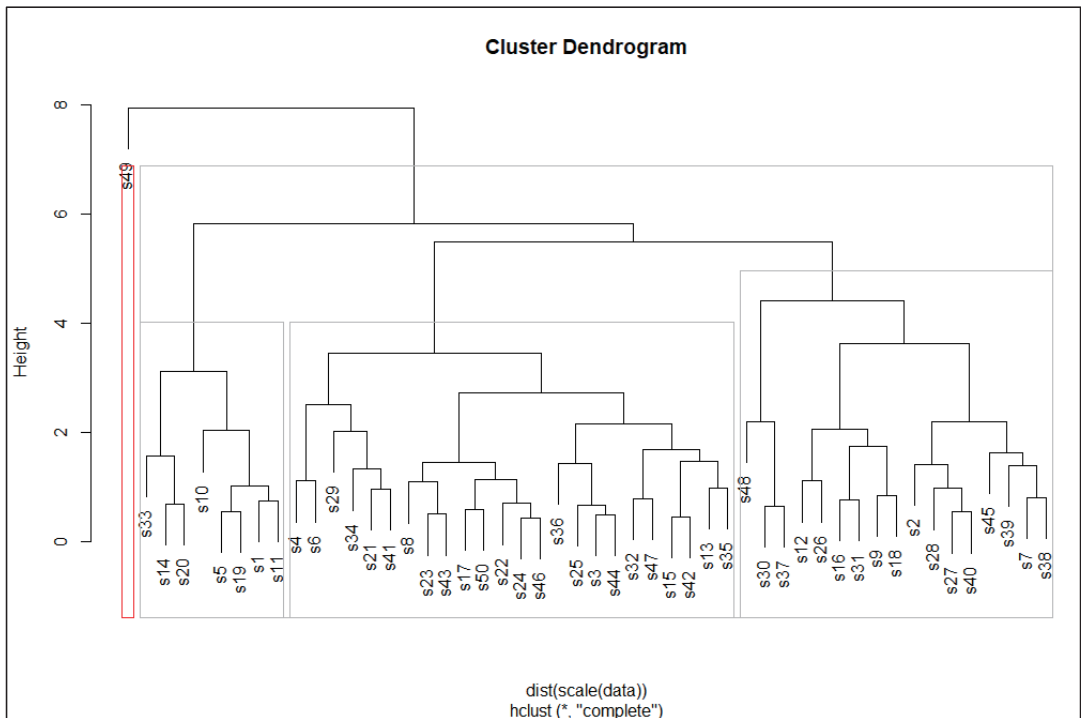
箱ひげ図から、mean2の方が箱の高さが小さくなっていることが分かる。またsdについても、sd2の方が大きくなっている。50人で評価したものがmean2であり、sd2であるので、アンケート調査で程良い標本平均やSDが得られるように、うまく評価点が出ている。mean1は各人が各

人の感覚で評価したものであり、個人差が大きく表れることが mean1 の箱が大きいことから分かる。このため、sd1には大きく離れた外れ値も多数出現している。

2. クラスタ分析

次に、今回のデータを用いて学生のタイプを分類することにした。解析には R4.2.1 を使用した。今回は、mean1、sd1、mean2、sd2 の 4 つを 1 組のデータとして扱った。

今回は距離の計算にはユークリッド距離を用いて、クラスタの結合には鎖効果を避けるために最遠隣法を使用した。結果のデンドログラムを掲載する。



結果を述べる。最遠隣法では 4 つのクラスターに分けられた。

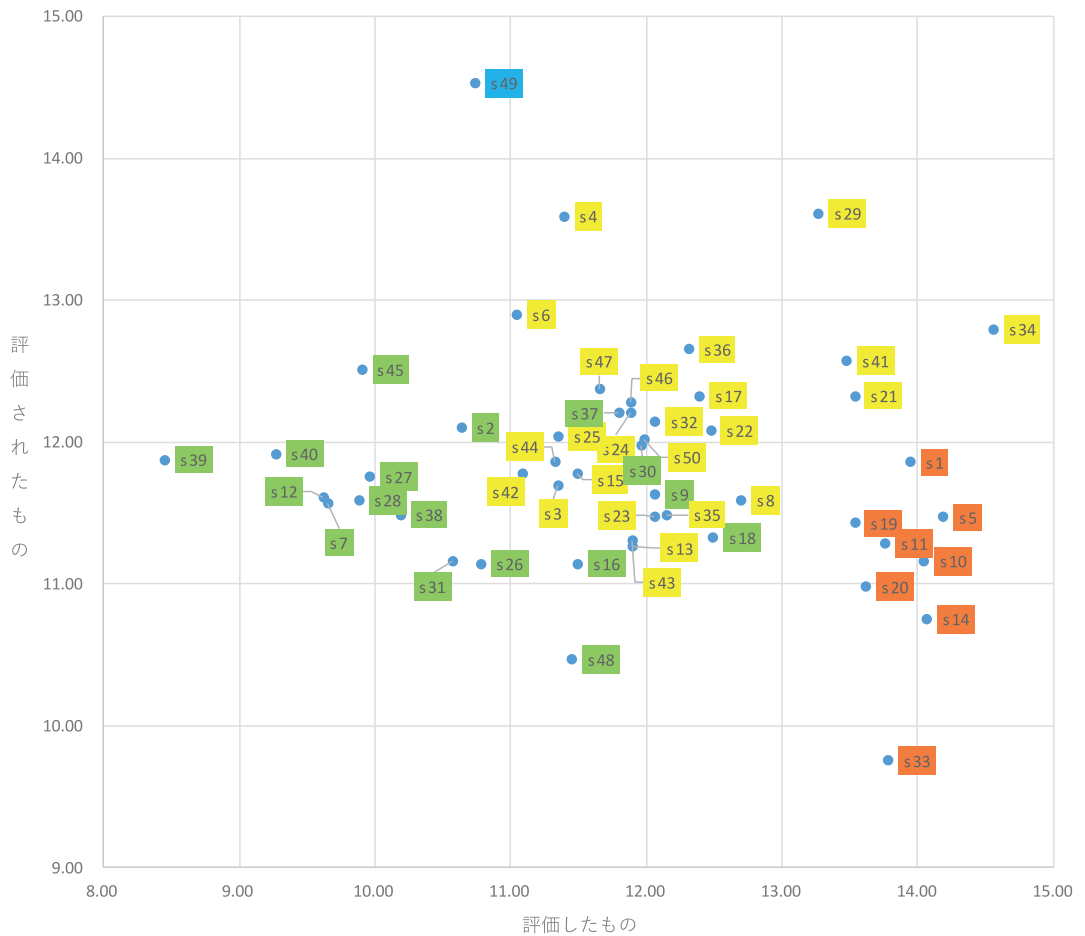
- ・ クラスター 1 : s1, s5, s10, s11, s14, s19, s20, s33 (8)
- ・ クラスター 2 : s3, s4, s6, s8, s13, s15, s17, s21, s22, s23, s24, s25, s29, s32, s34, s35, s36, s41, s42, s43, s44, s46, s47, s50 (24)
- ・ クラスター 3 : s2, s7, s9, s12, s16, s18, s26, s27, s28, s30, s31, s37, s38, s39, s40, s45, s48 (17)
- ・ クラスター 4 : s49 (1)

各クラスターの平均値を計算した。

	mean1	sd1	mean2	sd2
クラスター1	13.862	0.9680	11.087	1.9433
クラスター2	12.131	1.2316	12.173	1.7460
クラスター3	10.595	1.7128	11.617	1.9617
クラスター4	10.735	1.3367	14.531	0.9918

散布図上で各点のラベルに、クラスター 1 に赤、クラスター 2 に黄、クラスター 3 に緑、クラスター 4 に青を着色した。

評価した×評価された平均の分布



各クラスターの性格付けを行った。

- ・クラスター 1 (赤)：評価する点数は高いが評価される点数は低いグループ、他者には寛容な者。
- ・クラスター 2 (黄)：評価する・評価される点数ともに高いグループ。気前が良い者。
- ・クラスター 3 (緑)：評価する・評価される点数ともに低いグループ。自他共に厳しい者。
- ・クラスター 4 (青)：評価する点数は低いが評価される点数は高いグループ、他者に厳しい者。

クラスター 2 とクラスター 3 の境界部分は sd1 と sd2 の影響で振り分けられたと考えられる。

3. クラスターと GPA との関連

2節で行ったクラスター分けを活用して、各学生の GPA との関連を調べた。

今回の相互評価データは 3 年前学期の開講科目のため、3 年前学期までの累積 GPA のデータを各学生に割り当て、一要因分散分析を実施した。その結果は次のようになった。

まず各クラスターの群平均と全体平均は次のとおりである。

	クラスター1	クラスター2	クラスター3	クラスター4
群平均	2.5613	2.9063	2.6241	3.5500
全平均	2.7680			

次に、一要因分散分析の結果は次のとおりである。

	平方和	自由度	平均平方	F値	P
級間	1.7641	3	0.5880	3.9369	0.0139
級内	6.8709	46	0.1494		
全体	8.6350	49			

結果は 5% で有意であった。しかし、Tukey の HSD による多重比較の結果は各クラスター間での有意な差は見られなかった。

クラスターごとの GPA の平均値の大きい順は、

クラスター 4 > クラスター 2 > クラスター 3 > クラスター 1

の順であり、クラスター 1 がやはり他者評価は高いが自分の作品に対する評価は低いことが分かった。

4. 結果と考察

今回は学生の行った相互評価結果からクラスター分析を使用して、学生の評価状況をグルーピングした。本来は煩雑な計算処理が R4.2.1 を使用して比較的簡単に処理を行うことができた。また、遠距離法によって極めて効率的にグルーピングできたことは、今後の授業での活用の可能性が出てきたと思われる。

また、今回の解析を行って、学生同士による相互評価の結果を利用して、客観的に成績評価ができることが検証できたと思う。これまでも何回か学生による相互評価の結果を成績処理に活用していたが、今回の研究によって、健全な学生集団であれば実施の効果は十分にあることが分かった。

また、本論文の素案を本学特任教授の島治伸先生にお見せしたところ、3節で実施した「GPA との関係を見てはどうか」という知見を頂いた。実際に分析をした結果、分散分析は5%有意となり、相互評価結果をクラスターに分ける目的が得られた。島先生の有効な助言に感謝したい。

参考文献

兼子毅、R で学ぶ多変量解析、日科技連、2011.

赤間世紀、R リファレンスブック、カットシステム、2011.

上村龍太郎・北島良三・竹内晴彦・山下俊恵・吉岡茂、明日からビジネスで使える！Excel と R によるデータ解析入門、丸善出版、2014.