

筑波大学での倉谷賞候補選出方法

繁野 麻衣子, 山本 芳嗣

筑波大学社会学類同窓会が優秀な卒業研究に授与している倉谷賞への推薦学生の選出方法を報告します。社会学類の経営工学主専攻では卒業研究の発表に対して教員が表明した評価を Majority Judgment によって総合評価に集約して、その総合評価に成績を加味して推薦学生の候補を選んでいます。この方法に至った経緯や実施上の問題点を報告します。

キーワード：候補者選出, Majority Judgment, 評価語彙, 集約関数

1. はじめに

初代の社会学類長 倉谷好郎名誉教授のご寄付によって2006年に始まった倉谷賞は、社会学類の3主専攻、社会経済システム、経営工学、都市計画、のそれぞれからその年度の最も優秀な卒業生に授与されてきました。経営工学主専攻では、40~60件程度の卒業研究の発表を15~20人の教員が評価し、その年度の最も優れた卒業研究を倉谷賞の候補として同窓会組織である社工会に推薦することを当初から行ってきました。発表を重視してきたのは、倉谷賞受賞者には翌年度の社工会総会で先輩諸氏の前で発表することが義務付けられていることがその理由です。本稿ではその選考方法についてご紹介しようと思います。なお、倉谷賞と社工会については囲み記事や参考文献 [1] のサイトをご覧ください。

筑波社工会について

筑波社工会は、“筑波の地で、社会学を志したものの集まり”を目指した会です。筑波大学・社会学類の卒業生が母体ですが、理念は広く開かれた組織を目指しております。毎年、総会（筑波・東京を隔年）を開催し、その席でその年の倉谷賞の受賞者が社会学の勉強・研究の成果を発表して、社会学発展のサイクルの一助としております。

しげの まいこ

筑波大学システム情報系

〒305-8573 茨城県つくば市天王台 1-1-1

maiko@sk.tsukuba.ac.jp

やまもと よしつぐ

静岡大学大学院工学領域

〒432-8561 静岡県浜松市中区城北 3-5-1

yamamoto@sk.tsukuba.ac.jp

倉谷賞について

筑波大学理工学群社会学類の卒業生を対象に、“社会学を学んだ優秀な学生を励ましたい”との想いで、故 倉谷名誉教授のご寄付により、2006年に設立されました。社会学類は、社会システム・経営工学・都市計画の3主専攻に分かれており、原則として、毎年各専攻1名が受賞しています。受賞者の選考は、社会学類の教育を担当する教員会議で行われております。倉谷賞には、副賞として金一封が贈られており、筑波社工会が実務を担当しております。倉谷先生が永眠された後は、アメリカ在住の先生のお嬢様が志を受け継いで、副賞のご寄付を継続されています。

2. 当初の選考方法

2.1 有向グラフ上のゲーム

初年度は、各教員が発表を推薦に値すると考える順に順序付けました。各発表に対応させてグラフの頂点を作り、発表Aを発表Bよりも上位にランク付けした教員の人数を重みとする向きのある枝を、発表Aに対応する頂点から発表Bに対応する頂点につなげることによって重み付き有向グラフを作り、それをもとに発表をプレイヤーとするゲームを構成し、シャープレ値に類似したゲームの解を数種類計算し、各発表の評価としました。この年はこの評価で上位であった2名に優劣をつけがたく、結局2名を推薦することになりました。しかしこの方法は、発表を順序付けるのが大変であること、ゲームの解の計算が面倒などの実用上の欠点がありました。

2.2 得点の配分

次年度は前年度の反省を踏まえて、各教員が一律の

持ち点を良いと思った発表に配分し、配分された得点合計で発表に順序を付けるという方法を採用しました。発表が2日にわたって行われるため、初日に配分した点数を2日目に修正する必要がある、また良いと思う発表件数が増えたときに各発表に配分する点数が小さくなるため、自分の意見を最終結果に反映しづらくなるという印象でした。

3. Majority Judgment で

慣例として卒業研究の発表の世話は4年生のクラス担任の仕事です。筆者の1人がこの年の担任から相談を受け、大学近くのカフェで軽食を取りながら Majority Judgment を使ってみてはどうかと提案しました。

Majority Judgment はそもそもフランスのワイン品評会で用いられていた方法だそうですが、Balinski & Laraki がその著書 [2] やそれに先立つ何篇かの論文で提案した評価の方法です。詳しいことは本誌の解説 [3] やサイト [4] をご覧ください。本稿では評価される対象は発表で、評価する側は教員です。Majority Judgment では、各教員はあらかじめ決められた評価語彙にある評価語を用いて個々の発表に対する評価を表明します。評価語彙はたとえば $\{D, C, B, A\}$ や $\{0, 1, \dots, 100\}$ で、評価語の間には $D \prec C \prec B \prec A$ のように順序があります。

個々の発表に複数の教員が評価語を用いて与えた評価を何らかの方法で一つの評価語で表される総合評価に集約し、総合評価を比較することによって発表に順序を付け、最も好ましい発表を選びます。

3.1 評価語彙

Balinski & Laraki は評価語彙を共通言語と呼んでおり、評価語の意味が共有されていることが望ましいとし、フランスで小学校以来の成績評価に使われている言葉を選んでいきます。経営工学専攻では評価語彙として「該当せず、不十分、容認できる、まずまず、良い、非常に良い」を採用しました。

3.2 集約方法

Balinski & Laraki は一つの発表に対して表明された複数の評価をその発表の総合評価に集約する方法に、匿名性、全員一致性、単調性、戦略的操作耐性を要請して、そのすべてを満たす集約方法は以下のものしかないことを示しています。 k -順位関数とよばれているその集約方法は

1. 教員が提出した n 個の評価を昇順に並べたリストを作り、
2. そのリストの k 番目に位置する評価を総合評価と

する

というものです。ここで n は教員人数です。

Balinski & Laraki は k として中央に位置する評価（教員が偶数人の場合には中央に位置する二つの評価の低いほう）を選ぶことを薦めています。この重要なはずの k の値は前掲の4つの要請からは決めることができません。さらに、たとえば評価語彙が $\{D, C, B, A\}$ で、教員のほぼ半数が評価 D を、残りが評価 A を表明した場合にも順位関数は総合評価として D あるいは A を返すことになり、決して C や B を返しません。つまり評価が割れた場合に折衷的な評価を総合評価とすることがないのです。以上の問題点の解消方法については解説 [5] を参照してください。経営工学専攻では以上の問題点を承知したうえで、教員の評価が大きく割れることはないだろうと予想されたので、Balinski & Laraki の薦める順位関数をそのまま使うことにしました。

順位関数の例を見ておきます。教員の人数 n は10名、評価語彙は $\{D, C, B, A\}$ とします。たとえば教員の評価が $(C, C, A, B, A, C, D, C, B, A)$ なら、昇順に並べ直した

$$(D, C, C, C, \underline{C}, \underline{B}, B, A, A, A) \quad (1)$$

の中央に位置する $\underline{C}, \underline{B}$ の低いほうの評価である C が総合評価となります。総合評価よりも高い評価である B を付けていた教員は自分の評価を修正して総合評価を引き上げる動機をもちますが、仮にその評価を A に引き上げても総合評価を変更することができないことを確認してください。これが戦略的操作耐性です。一方、 D を付けていた教員がその評価を B に変えたとすると $(C, C, C, C, \underline{B}, \underline{B}, B, A, A, A)$ の中央に位置するのは $\underline{B}, \underline{B}$ ですから B が新たな総合評価となります。しかし、この教員は総合評価 C よりも低い評価である D を付けていたのですから、評価を B に変更して総合評価を上げる動機をもたないことに注意してください。

3.3 同点の解消

発表件数は評価語の個数よりも多いため、以上のように決まる総合評価では同点が起こります。たとえば上の (1) も

$$(D, D, D, C, \underline{C}, \underline{C}, C, B, B, B) \quad (2)$$

も総合評価は C です。この両者を区別するために総合評価より低い評価と高い評価の割合を用います。

$$(\underline{D}, C, C, C, C, \underline{B}, \underline{B}, A, A, A) \rightarrow (0.10, 0.50)$$

$$(\underline{D}, \underline{D}, \underline{D}, C, C, C, \underline{C}, \underline{B}, \underline{B}, B) \rightarrow (0.30, 0.30)$$

卒研発表会評価												成績
No.	5 非常に 良い	4 良い	3 まあま あ	2 容認で きる	1 不十分	0 該当せ ず	平均	中央値	高評 価割 合	低評 価割 合	評価者 人数	
1	4	3	1	0	0	0	4.38	4	0.50	0.20	8	
2	3	7	1	0	0	0	4.18	4	0.27	0.09	11	
3	0	4	3	0	0	0	3.57	4	0.00	0.43	7	
4	1	3	5	0	0	0	3.56	3	0.44	0.00	9	
5	1	4	2	1	0	0	3.38	4	0.13	0.38	8	
6	0	4	3	1	0	0	3.13	3	0.50	0.13	8	
7	1	2	4	1	0	0	3.13	3	0.40	0.13	8	

図 1 評価の集計抜粋

当然、前の数字が小さく、後ろの数字が大きいほうがより高く評価されることになります。

4. 実施上の問題点と変更点

2日にわたって行われる卒業研究の発表会の日には通常の授業が行われているうえ、教員の何人かが出席しなければならない会議が入ったりします。そのため発表を聞く教員の人数が一定しません。10人以上の教員から評価を受けている発表がある一方で、数人の教員からしか評価を受けていない発表があったりします。この点が総合評価の信頼性を低下させることになっていました。

そのために数年前から教員と発表を三つのグループに分け、教員は割り当てられたグループの発表のすべてに出席する方式に変更しました。割り当てられなかったグループの発表への出席は任意です。各グループで、Majority Judgmentを使って総合評価を決め、最も高い総合評価を得た発表1件をそれぞれのグループから推薦し、推薦された3件の発表について、その学生の成績を考慮して、経営工学主専攻からの最終的な推薦学生を決めることにしました。2016年度の卒業研究発表のあるグループの評価を集計した一部分が図1です。行が発表に対応し、評価語の列に書かれた数字はその評価をした教員の人数です。また評価を0から5に数値化してその平均の順に並べてあります。番号4と5の発表について平均と中央値が逆転していることがわかります。また中央値が4である発表（網掛けの行）について、高評価割合と低評価割合を見ると、番号1と2の発表が番号3と5よりも優れていることが読み取れます。

5. おわりに

最も優れた卒業研究が受賞候補となることが、よい卒業研究、よい発表への学生たちの動機づけにつながっていると思われます。そのために公平な評価方法が求められますが、本方法に対してこれまで教員と学生のいずれからも不満の声が上がったことはありません。これは、学生の発表に対する教員の評価が大きく割れることがなく、また選ばれた発表が飛び抜けて素晴らしいこともその一因です。一方で、評価方法に対しては理論の後ろ盾があることから、教員も自信をもって候補者を推薦できています。全教員の出席が要望されていながら実現していないという問題点は残されていますが、Majority Judgmentはうまく機能していると思われます。

謝辞 囲み記事を書いてくださった社工会会長の杉山幸登氏に感謝します。

参考文献

- [1] <https://www.sk.tsukuba.ac.jp/College/outline/kuratani.html> (2017年11月8日閲覧)
- [2] M. Balinski and R. Laraki, *Majority Judgment: Measuring, Ranking, and Electing*, MIT Press, 2010.
- [3] 山本芳嗣, “1人1票から Majority Judgmentへ,” オペレーションズ・リサーチ: 経営の科学, **57**(6), pp. 295–301, 2012.
- [4] http://infoshako.sk.tsukuba.ac.jp/~yamamoto/Majority_Judgment/1._Majority_Judgment_heyokoso.html (2017年11月8日閲覧)
- [5] 山本芳嗣, “複数の評価の集約方法—割れた評価をまとめれば上手くいく—,” オペレーションズ・リサーチ: 経営の科学, **62**(10), pp. 647–659, 2017.