

[原著論文]

クラスサイズが授業評価に与える影響について

入江 雅仁*

About the impact of class size on class evaluation

Masahito IRIE*

キーワード：授業評価アンケート，クラスサイズ，基本統計量，ウェルチの検定

*九州共立大学経済学部

*Faculty of Economics, Kyushu Kyoritsu University

1. はじめに

文部科学省(2022)³⁾によると、学生による授業評価(以下では、授業評価と略記)は、令和2年度において、国内の98.7%の大学で導入されており、また、授業評価に関する特徴的な取り組みとして、授業評価アンケートの結果を組織的に検討し、授業内容等に反映する機会を設けている大学は71.4%にのぼっている。

九州共立大学も、学是「自律処行」に基づく三つ(AP・CP・DP)の方針の下で実行している教育活動の点検・評価の一環として、授業評価アンケートを実施し、その集計結果を担当教員へフィードバックしている¹⁾。

しかしながら、授業評価はさまざまな要因²⁾の影響を受けているにもかかわらず、担当教員へフィードバックされる情報は各科目の評価点と各設問の全学平均点などの記述統計の情報にとどまるだけでなく、授業評価の分析は各担当教員の裁量に任されており、それらの要因が授業評価に与える影響についての全学的な検証が残されたままとなっている。

そこで本論では、授業評価の全学的な検証の前段階として、クラスサイズ(受講者数)の観点から一教員の授業評価に関する統計的仮説検定を実施する。具体的には、受講者数が100名以上の大規模クラスと100名未満の一般クラスに分類したときに、クラスサイズによって一教員の授業評価に統計的な有意差があるかどうかをウェルチのt検定で検証する。

なお、クラスサイズに着目した理由は、以下の3点である。一つ目は、大規模クラスを担当する場合、授業運営(授業準備・出席管理・採点作業など)に多大な労力がかかるので、教員のこうした労力を評価すべきと考えていたからである。二つ目は、大規模クラスを担当する場合、一般クラスより選択可能な教授法の自由度が制限されるので、学生の理解度を高めるための創意工夫や教員の試行錯誤を評価すべきと考えていたからである。三つめは、大規模クラスを担当する場合、「授業評価の観点で不利であると考えられる(南(2004, p.59)²⁾)」にも関わらず、クラスサイズを考慮しない同一の全学平均点しかフィードバックされていない点に疑問を感じていたからである。

¹⁾ 各担当教員は、授業評価を分析・省察し、その内容を次年度のシラバスへ反映させ、授業改善へ繋げている。

²⁾ 例えば、対象科目の学問領域(カリキュラムにおける科目の位置づけ)、授業形態やクラスサイズ(受講者数)をはじめとした教育環境、教育方法や授業運営能力をはじめとした教員の特性、教員の先天的な資質、学生の受講態度や専攻分野をはじめとした学生の特性などが考えられる。

2. 調査方法

1) 調査対象科目

授業評価アンケートの調査対象は、2012年度から2022年度の間に筆者が担当した科目で、具体的な科目名は付録のTable3の通りである。なお、2012年度から2015年度までは非常勤講師として、2016年度から専任講師として担当した科目である。ただし、新型コロナウイルスの影響で簡易的な授業評価アンケートとなった2020年度前期を分析対象外とした点と³⁾、本論では全学的な検証ではなく一教員の授業を対象としているため、得られたデータにおいて教員の多様性を持たない点に注意が必要である。

2) 調査内容

本論で用いた授業評価アンケートの設問内容(調査内容)は付録のTable4の通りである。なお、このTable4における設問分類は、設問番号 $k = 1, \dots, 4$ を学生の受講態度に関する設問、 $k = 5, 9$ を教員による授業運営に関する設問、 $k = 7, 8$ を教員の(先天的な)資質に関する設問、 $k = 10, 11, 13, 14$ を教育方法に関する設問、 $k = 6, 12, 15$ を科目特性に関する設問として筆者が独自に分類したものである。

また、各設問に対する回答の選択肢は5段階評価で調査を行い、Table4の①に1点、②に2点、③に3点、④に4点、⑤に5点を付与して算出する科目 $i (= 1, \dots, n)$ に関する設問項目 $k (= 1, \dots, 15)$ の評価 X_k^i について分析する：

$$X_k^i = \frac{\sum_{a=1}^5 \text{設問項目 } k \text{ の } \textcircled{a} \text{ の数} \times a}{\sum_{a=1}^5 \text{設問項目 } k \text{ の } \textcircled{a} \text{ の数}}.$$

なお、記号 n は調査対象科目の総科目数である。

3) 調査データの基本統計量

本節では、統計的仮説検定に先立って、設問項目 $k (= 1, \dots, 15)$ の基本統計量(平均 \bar{x}_k^\bullet 、不偏標準偏差 U_k^\bullet 、歪度 $\beta_{1,k}$ 、および、尖度 $\beta_{2,k}$)を確認する。

³⁾ 2020年度前期の対象科目は演習科目のみで、かつ、設問項目も限定的であったため、対象外とした。

$$\begin{aligned}\bar{X}_k^* &= \frac{1}{n} \sum_{i=1}^n X_k^i, \\ U_k^* &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_k^i - \bar{X}_k^*)^2}, \\ \beta_{1,k} &= \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{X_k^i - \bar{X}_k^*}{U_k^*} \right)^3, \\ \beta_{2,k} &= \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{X_k^i - \bar{X}_k^*}{U_k^*} \right)^4 \\ &\quad - \frac{3(n-1)^2}{(n-2)(n-3)}.\end{aligned}$$

上記の実現値をExcelの関数⁴で求めた結果はTable1の通りである。なお、次節以降、設問項目 k の評価について正規分布を仮定する。

Table 1 設問の集計結果 ($n=100$)

k 設問項目	平均 \bar{x}_k	不偏標準 偏差 u_k	歪度 $\beta_{1,k}$	尖度 $\beta_{2,k}$
1 出席回数	4.05	0.42	-0.79	1.60
2 参加意欲	4.04	0.33	0.24	1.80
3 シラバス閲覧	3.65	0.50	-1.43	7.71
4 予復習	3.55	0.42	0.64	1.57
5 注意喚起	4.15	0.31	0.20	2.20
6 授業難易度	3.91	0.46	-0.18	0.83
7 聞きやすさ	4.02	0.41	-0.11	0.40
8 熱意	4.04	0.38	-0.12	1.29
9 時間管理	4.17	0.32	0.29	0.16
10 目標の明示度	4.05	0.37	-0.32	1.44
11 参加の促し	4.01	0.51	-0.37	0.34
12 授業理解度	3.86	0.48	-0.18	0.31
13 教材の適正さ	4.03	0.39	-0.01	0.53
14 見やすさ	3.96	0.45	-0.34	0.52
15 満足度	3.97	0.47	-0.28	0.45

3. ウェルチの検定

ここでは、調査対象科目 ($n=100$) を、受講者数が100名以上の大規模クラス群Aと受講者数が100名未満の一般クラス群Bに分け、両群の授業評価アンケートにおける設問項目 k の評価の母平均に違いがあるかどうかについて、ウェルチの検定を有意水準5%で行う。

すなわち、大規模クラス群Aおよび一般クラス群Bの設問項目 k の母平均をそれぞれ、 μ_k^A および μ_k^B ($k=1, \dots, 15$) とし、設問項目 k の帰無仮説 $H_0: \mu_k^A = \mu_k^B$ 、対立仮説 $H_0: \mu_k^A \neq \mu_k^B$ について、両側検定を行う。なお、以下では大規模クラス群のサンプルサイズを n^A 、標本平均を $\bar{X}_k^A = \frac{1}{n^A} \sum_{i \in A} X_k^i$ 、不偏標本標準偏差を $U_k^A = \sqrt{\frac{1}{n^A-1} \sum_{i \in A} (X_k^i - \bar{X}_k^A)^2}$ 、同様に、一般クラス群のサンプルサイズを n^B 、標本平均を $\bar{X}_k^B = \frac{1}{n^B} \sum_{i \in B} X_k^i$ 、不偏標本標準偏差を $U_k^B = \sqrt{\frac{1}{n^B-1} \sum_{i \in B} (X_k^i - \bar{X}_k^B)^2}$ とする⁵。

Table 2 ウェルチの検定統計量 T_k の実現値

設問項目	クラス サイズ	科目数 n^i	平均値 \bar{x}_k^i	不偏標準偏差 u_k^i	検定統計量 t_k	自由度 ν_k
1 出席回数	大規模	23	4.16	0.01	2.43*	96 \approx 96.2
	一般	77	4.01	0.22		
2 参加意欲	大規模	23	3.91	0.03	-3.00**	74 \approx 73.7
	一般	77	4.08	0.12		
3 シラバス閲覧	大規模	23	3.67	0.05	0.44	88 \approx 88.3
	一般	77	3.64	0.30		
4 予復習	大規模	23	3.51	0.06	-0.73	67 \approx 67.2
	一般	77	3.56	0.21		
5 注意喚起	大規模	23	4.05	0.02	-2.64**	84 \approx 84.1
	一般	77	4.18	0.11		
6 授業難易度	大規模	23	3.59	0.06	-5.61**	70 \approx 70.1
	一般	77	4.01	0.22		
7 聞きやすさ	大規模	23	3.76	0.06	-5.08**	65 \approx 64.5
	一般	77	4.10	0.17		
8 熱意	大規模	23	3.80	0.04	-4.98**	74 \approx 74.1
	一般	77	4.11	0.15		
9 時間管理	大規模	23	4.04	0.03	-3.37**	76 \approx 75.7
	一般	77	4.21	0.12		
10 目標の明示度	大規模	23	3.86	0.04	-4.03**	71 \approx 71.0
	一般	77	4.11	0.15		
11 参加の促し	大規模	23	3.62	0.05	-6.74**	84 \approx 84.1
	一般	77	4.13	0.26		
12 授業理解度	大規模	23	3.51	0.07	-5.75**	66 \approx 66.3
	一般	77	3.97	0.24		
13 教材の適正さ	大規模	23	3.79	0.05	-4.77**	70 \approx 69.6
	一般	77	4.10	0.16		
14 見やすさ	大規模	23	3.67	0.07	-4.90**	63 \approx 62.6
	一般	77	4.05	0.21		
15 満足度	大規模	23	3.65	0.06	-5.48**	71 \approx 71.2
	一般	77	4.06	0.23		

* $p < 0.05$, ** $p < 0.01$

t 分布の自由度は小数点第1位を四捨五入した整数を採用した。

芝ら(1999, p.13)⁵⁾によると、ウェルチの検定で用いる以下の検定統計量 t_k は、設問項目 k の評価が非等分散の正規分布に従うと仮定すると、帰無仮説 H_0 のもとで、自由度 ν_k の t 分布に近似的に従う：

$$t_k = \frac{\bar{X}_k^A - \bar{X}_k^B}{\sqrt{\frac{U_k^{A^2}}{n^A} + \frac{U_k^{B^2}}{n^B}}} \sim t_k(\nu),$$

ただし、自由度 ν_k は次の右辺を小数点第1位で四捨五入して求めた：

$$\nu_k \approx \frac{\left(\frac{U_k^{A^2}}{n^A} + \frac{U_k^{B^2}}{n^B} \right)^2}{\frac{U_k^{A^4}}{n^A{}^2(n^A-1)} + \frac{U_k^{B^4}}{n^B{}^2(n^B-1)}}.$$

Excelの分析ツール「 t 検定：分散が等しくないと仮定した2標本による検定」から求めた実現値はTable2に示されている。そのTable2に示された設問項目 k の検定統計量 t_k と、自由度 ν_k 、有意水準 5%、および、 t 分布表から、棄却条件 $t_k \leq t_{\nu_k}(0.05/2)$ の不成立・成立に応じて帰無仮説の採択・棄却を判定した。

⁴ ここでは、設問項目 k の平均を「AVERAGE」関数、不偏標準偏差を「STDEV」関数、歪度を「SKEW」関数、尖度を「KURT」関数で求めた。

⁵ 記号 $\sum_{i \in A} X_k^i$ は、大規模クラス群に該当する科目 i の設問項目 k の評価を集計した値であり、同様に、記号 $\sum_{i \in B} X_k^i$ は、一般クラス群に該当する科目 i の設問項目 k の評価を集計した値である。

4. 考察

1) 基本統計量に関する考察

Table1から次の点が指摘できるであろう。

受講態度 出席回数 ($k = 1$) や参加意欲 ($k = 2$) が平均4を超えている一方で、予復習 ($k = 4$) が平均3.55 と最小となっているので、学生は意欲的に授業へ出席しているにもかかわらず、予復習という具体的な学習行動に結びついていない可能性⁶⁾が示された。また、シラバス閲覧 ($k = 3$) が平均 3.65 と 2 番目に低い値となっていることから、学生はシラバスに対する認識が低く、シラバスを重要視していない可能性が明らかになった。

科目特性 難易度 ($k = 6$) ・理解度 ($k = 12$) ・満足度 ($k = 15$) がいずれも平均4を下回っていることから、学生にとって難しい科目を担当している可能性が確認できた。なお、担当科目の特性上、数式やグラフなどを使った授業を展開することが多い点と授業評価に関係があるのではないかと考えているので、今後の研究でこの点を検討したいと考えている。

歪度 参加意欲 ($k = 2$) ・予復習 ($k = 4$) ・注意喚起 ($k = 5$) ・時間管理 ($k = 9$) は正の歪度なので、右の裾が長く左に偏った（相対的に高評価が少なく、逆に低評価が多い）分布を持つことから、学生は辛口の評価を下している可能性が明らかになった。他方で、それら以外の設問項目 ($k = 1, 3, 6, 7, 8, 10, 11, 12, 13, 14, 15$) は負の歪度なので、左の裾が長く右に偏った（相対的に低評価が少なく、逆に高評価が多い）分布をしていることから、学生は甘口の評価を下している可能性が明らかになった。

2) ウェルチの検定に関する考察

Table2から次の点が指摘できるであろう。

シラバス閲覧 ($k = 3$) と予復習 ($k = 4$) を除く13（約87%）の設問項目 ($k = 1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$) で、大規模クラス群と一般クラス群の評価に有意差が認められた。この結果は、「クラス規模が授業満足度に部分的にでも影響があること」(p.41) を指摘した荒木 (2016)¹⁾ の結果を包括したものと解釈できる⁷⁾。また、年度や開講期の違いに注目した田実・竹原 (2008)

⁶⁾ でも設問項目の約80%⁸⁾ で有意差が認められており、本研究と整合的な結果が得られたのではないかと考えている。なお、ほとんどの設問項目（約87%）で有意差が認められたので、クラスサイズに応じた授業改善を進めることは有益だと推察される。したがって、授業改善を進め、満足度を向上させるためには、クラスサイズに応じた情報を担当教員にフィードバックすることが必要である。

また、受講生が少ない一般クラス群Bの授業理解度 ($k = 12$) や満足度 ($k = 15$) の評価が大規模クラス群Aのそれより高い $\bar{X}_k^A < \bar{X}_k^B$ ことから、クラスサイズが小さな科目を担当することが多くなるほど、それらの評価が相対的に高くなっていることが予想される。なお、もし一教員ではなく、全学的にクラスサイズで授業評価に偏り（有意差）が生じるならば、授業評価に基づいて翌年度の受講制限やクラス分割を設定するなどの活用法も考えられる。

他方、受講者数が多い大規模クラス群Aの評価（平均値）が一般クラス群Bより低くなると予想していたけれども、出席回数 ($k = 1$) とシラバスの閲覧 ($k = 3$) は大規模クラス群の評価が高い $\bar{X}_k^A > \bar{X}_k^B$ ことから、大規模クラス群が授業評価の観点から一概に不利と言えないことを確認できた。なお、統計的な有意差が見出されなかったシラバスの閲覧 ($k = 3$) で、そのような結果になった要因として、講義科目でシラバスの授業計画を毎週提示している一方で、一般クラスの半数以上⁹⁾を占める演習科目でシラバスの授業計画を毎週提示していないことが関係していると考えられる。

5. おわりに

本論では、一教員の授業評価アンケートを対象に、各設問項目に関する基本統計量を確認し、以下の点などがTable 1 から明らかになった。

- 学生は意欲的に授業へ出席しているにもかかわらず、予復習という具体的な学習行動に結びついていない。
- 学生はシラバスに対する認識が低く、シラバス

⁷⁾ 荒木 (2016)¹⁾ は本論における科目特性の満足度 ($k = 15$) に対応する架空データでクロス集計表を作成しているものの、その他の項目については作成していない点に注意が必要である。

⁸⁾ 田実・竹原 (2008)⁶⁾ では、13項目中10項目で2003年度と2005年度の評価に有意差が認められ、また、2003年度の前期と後期の評価は13項目中11項目で、他方、2005年度の前期と後期の評価は13項目中11項目で有意差が認められた。

⁹⁾ 本論で一般クラスに分類した 77 科目のうち、48 科目が演習科目である。

⁶⁾ 森・田邊 (2010)⁴⁾ も、近畿大学の授業評価の結果から、「学生は教員が授業に取り組む姿勢を高く評価しているにもかかわらず、授業自体には興味を持てず、それゆえに実際の学習行動には結びついていない実態 (p.736)」を指摘している。

を重要視していない。

- 学生にとって難しい科目を担当している。
- 参加意欲 ($k = 2$)・予復習 ($k = 4$)・注意喚起 ($k = 5$)・時間管理 ($k = 9$) について、学生は辛口の評価を下している。
- 設問項目 ($k = 1, 3, 6, 7, 8, 10, 11, 12, 13, 14, 15$) について、学生は甘口の評価を下している。

また、大規模クラス群(受講者100名以上)と一般クラス群(受講者100名未満)で授業評価の母平均に統計的な有意差があるかをウェルチの検定で検証した Table 2 から、以下の点が明らかになった。

- ウェルチの検定で、約87%の項目で統計的な有意差が見出された。
- 大規模クラス群が授業評価の観点から一概に不利と言えない。

最後に今後の課題について言及する。まず文献調査を充実させ、授業評価に影響を与えている要因を整理する必要がある。次に、それらの要因を踏まえて、重回帰分析や共分散構造分析を実行し、満足度を高めるための具体的な方法を数量的に検討する必要がある。また、一教員だけでなく全学的な検証を行い、どのような要因が授業評価に影響を与えているかを検証する必要がある。

6. 謝辞

本研究は、九州共立大学の令和5年度 特別教育研究費(学長政策費)による助成を受けて実施した研究の成果である。このような研究の機会を与えていただいた九州共立大学奥田俊博学長に心よりお礼を申し上げます。

参考文献

- 1) 荒木俊博 (2016): 「履修者人数と授業アンケート結果の関連についての検討」, 『大学評価とIR』, pp.36-43.
- 2) 南学 (2004): 「学生による授業評価におけるクラスサイズの効果」, 松山大学総合研究所, 『松山大学論集』, 第16巻(2), pp.57-75.
- 3) 文部科学省 高等教育局(2022): 「令和2年度の大学における教育内容等の改革状況について(概要)」 < 3 - J 学生による授業評価の実施状況 >, https://www.mext.go.jp/content/20230117-mxt_daigakuc01-000025974_1r.pdf, 閲覧日: 2023/6/22.
- 4) 森節子, 田邊義隆 (2010): 「授業評価アンケート調査から読み取れる学生の意識と授業の課題-近畿大学法学部における現状分析」, 近畿大学法学会,

『近畿大学法学』, 第58巻(2/3), pp.721-742.

- 5) 芝祐順, 渡部洋, 石塚智一[編](1999), 『統計用語辞典』, 新曜社, p.13.
- 6) 田実潔, 竹原卓真 (2008): 「学生による授業評価に基づいた授業改善への探索的研究: 授業評価アンケートの分析から」, 北星学園大学, 『北星学園大学社会福祉学部北星論集』, 第45巻, pp. 37-43.

表付録

Table 3 調査対象科目

担当年度	職位	科目名	履修区分	科目区分	授業形態
2014-2016	非常勤	経済数学	専門	関連	講義
2012-2013	非常勤	ミクロ経済学入門	専門	基礎	講義
2013-2015	非常勤	ミクロ経済学	専門	コース	講義
2014-2015	非常勤	産業組織論入門	専門	コース	講義
2014-2015	非常勤	産業組織論	専門	コース	講義
2015-2016	専任	貿易論	専門	コース	講義
2016	専任	経済成長論	専門	コース	講義
2018	専任	演習 II	専門	演習	演習
2016	専任	演習 III	専門	演習	演習
2017	専任	演習 IV	専門	演習	演習
2017-2018	専任	演習 V	専門	演習	演習
2016-2022	専任	福原学	共通	コア	演習
2016-2022	専任	職業とコミュニケーション	共通	キャリア	演習
2017-2022	専任	マクロ経済学入門	専門	コア	講義
2017-2021	専任	マクロ経済学	専門	コア	講義
2017-2022	専任	人口学	専門	領域	講義
2016-2022	専任	公共経済学入門	専門	領域	講義
2016-2022	専任	公共経済学	専門	領域	講義
2016-2017	専任	経済統計	専門	領域	講義
2016-2022	専任	ゼミナール入門	専門	ゼミナール	演習
2016-2022	専任	ゼミナール I	専門	ゼミナール	演習
2017-2022	専任	ゼミナール II	専門	ゼミナール	演習
2018-2022	専任	ゼミナール III	専門	ゼミナール	演習
2019-2022	専任	ゼミナール IV	専門	ゼミナール	演習
2018-2022	専任	ゼミナール V	専門	ゼミナール	演習

Table 4 アンケート設問一覧表

設問番号	アンケート設問内容	設問項目	設問分類
1	あなたは、この授業にどの程度出席しましたか?	出席回数	受講態度
2	あなたは、意欲的にこの授業に臨みましたか?	参加意欲	受講態度
3	あなたは、この授業のシラバス(講義要項)を読みましたか?	シラバス閲覧	受講態度
4	あなたは、この授業について、予習・復習をどの程度行いましたか?	予復習	受講態度
5	教員は、私語や携帯電話を講義中に使用している学生に対して注意をし、授業の静粛を確保していましたか?	注意喚起(静粛確保)	授業運営
6	授業内容の難易度(レベル)は、あなたにとって適切でしたか?	授業難易度	科目特性
7	教員の説明は、明瞭で聞き取りやすかったですか?	聞きやすさ	教員の質
8	授業に対する教員の熱意を感じましたか?	熱意	教員の質
9	教員は、授業時間を守っていましたか?	時間管理	授業運営
10	教員は、この授業... 科目の学習目標をはっきり示しましたか?	目標の明示度	教育方法
11	教員は、一方的な授業ではなく、学生の参加(意見・発表)を促しましたか?	参加の促し	教育方法
12	授業内容は理解できましたか?	授業理解度	科目特性
13	授業で使っていた教科書・配布資料は適切でしたか?	教材の適切さ	教育方法
14	黒板・スライド等は見やすかったですか?	見やすさ	教育方法
15	この授業における満足度は如何でしたか?	授業満足度	科目特性

Received date 2023年7月26日

Accepted date 2023年7月26日