

授業評価(聴覚部・13年度)の統計解析

第1報 授業評価の原データの平均値とばらつきが語るもの

筑波技術短期大学 聴覚部¹⁾, 同 機械工学科²⁾, 同 一般教育等³⁾, 同 デザイン学科⁴⁾, 同 建築工学科⁵⁾,
同 電子情報学科電子工学専攻⁶⁾, 同 電子情報学科情報工学専攻⁷⁾, 同 教育方法開発センター⁸⁾

聴覚部 教育活動に関する点検評価 ワーキンググループ¹⁾

三牧敏太郎²⁾ 根本匡文³⁾ 生田目美紀⁴⁾ 萩田秋雄⁵⁾ 川島光郎⁶⁾

渡辺隆⁶⁾ 皆川洋喜⁷⁾ 松藤みどり³⁾ 石原保志⁸⁾ 中瀬浩一⁸⁾

要旨: 聴覚部の“教育活動に関する点検評価ワーキンググループ”は平成11~13年度にわたり、学生による授業評価アンケートの統計解析を継続している。本報では、アンケートの原データを用いた平均値の差の検定などを行い、アンケート項目の「評価」、「理解」と「興味」の平均値に関する次の結果を報告する。①「評価」、「理解」と「興味」は授業形式(講義系、演習系)による差があり、演習系の方が高い。②講義系の「評価」、「理解」と「興味」は学科による差はない。③演習系では、一般教育は高いが、他の5学科・専攻は有意差なしとして良い。

キーワード: 授業評価, アンケート調査, 平均値の差の検定, 分散分析

1. 緒言

聴覚部において、平成11年度から“学生による授業評価”のアンケート調査が実施されている。アンケートは付録-1に示す12項目からなり、それらは次のように分類できる。学生側の要因は「評価」、「理解」、「興味」、「態度」の4つ、学生-教官に共通する要因は「疎通」、「通話」、「受話」の3つ、教官側の要因は「目的」、「機器」、「準備」、「熱意」、「程度」の5つである。

これまでにアンケートデータの統計解析に基づいて、有用な報告がなされている^{1)~4)}。これらの報告に用いられている統計手法は、「評価」の平均値に対して差別化する主成分を導き、その分散を最大にする主成分分析¹⁾と「評価」などを特性値に他のアンケート項目を説明変数にし、特性値に対する関数式を導く回帰分析²⁾である。しかし、原データに関しては、その解析結果は報告されていない。ここで、原データ(以下で、データ)とは、各項目の5段階評点(1~5)を回答数で重み付けした平均値 μ のことであり、データ数(科目数) N は、 $N=76$ (H11年度)、 $N=105$ (H12年度)、 $N=110$ (H13年度)である。

本報では、アンケート調査の目的とも言える「評価」を取り上げ、「評価」の平均値に対して、数値化できない要因(層別因子)である“授業形式(講義系授業、演習系授業)と学科”の有差の検定結果を報告する。併せて、学生の知識獲得の目安としての「理解」、学生の勉学への心構えとしての「興味」に対する層別因子の検定結果も示す。また、「評価」と「理解」が全授業の平均値から大きく偏ったデータ(科目)を参照して、「評価」の高低に対する授業内容に関する仮説を導いたので、それを紹介する。

ここで、授業形式は授業計画(シラバス)を参照し分類しているが、演習系授業は演習、実験、実習、実技などの講義以外の授業であり、①講義で学んだことを実際に体験することで、

理解の定着や思考過程の育成を図る授業、②作品の制作や保健体育などの実技能力の育成・向上を図る授業 などである。

2. 「評価」のデータと解析手順

「評価」の全授業、授業形式、学科における各年度の平均値を表1に示す。ここで、全授業=(講義系+演習系)、全学科= Σ (学科)である。

各年度のデータ数 N で重み付けした3年間での平均値は、全授業 $\mu_{ALL}=3.94$ 、講義系 $\mu_{LEC}=3.80$ 、演習系 $\mu_{EXE}=4.14$ である。これより、授業形式による「評価」の差(演習系-講義系) $\mu_{DEF}=\mu_{EXE}-\mu_{LEC}=0.34$ 、 $\mu_{DEF}/\mu_{EXE}=8.2\%$ となる。

表1に示した各学科の「評価」の平均値(全授業)の年度変化を図1に示す。同図において、実印(ソリッド)の記号は一般教育(●)、デザイン(▲)と建築(■)を、空印(オープン)は機械(○)、電子(△)と情報(□)の各学科・専攻を表している。空実の2グル

表1 「評価」の各年度の平均値

「評価」の平均値	H11	H12	H13
全授業 (全学科)	3.86	3.97	3.97
講義系 (")	3.74	3.83	3.80
演習系 (")	4.01	4.16	4.22
一般教育 (全授業)	3.99	4.03	4.06
デザイン (")	4.18	4.35	4.23
機械 (")	3.99	3.84	3.90
建築 (")	3.93	4.12	3.97
電子 (")	3.68	3.85	3.62
情報 (")	3.71	3.60	3.87
データ数 全授業(全学科)	76	105	110
" 講義系(")	41	62	66
" 演習系(")	35	43	44

ープに区分したのは、空印で表す学科は『数学に関係する難しい授業が多くて、学生による「評価」も悪くなりがちである』という意見⁴⁾があるのに対応させるためである。

図1に見るように、実印のデータは空印のデータよりも「評価」は相対的に高く、「評価」はグループにより偏りがあるとの意見に符合している。しかし、以下に示す“学科間の平均値の差”の検定によれば、“「評価」は学科による差はほとんどない”と99%に近い確率で言える(3.2節参照)。したがって、前述の意見は6学科を2グループに層別した場合におけるものと言える。

「評価」の平均値の“偏り=差の有無”を判定するには、2つの平均値の差を両側検定(差の土での有意差の検定)すればよい。小標本から求められた平均値はスチューデントの t 分布することが知られており、平均値の差の検定法もいくつかある。検定法の概略を紹介する。データから t_0 で表す t 値と自由度 ϕ が求められる。また、判定の過誤の危険率 α を設定すれば、 t 表より $t(\phi, \alpha)$ の値が与えられる。 $|t_0| > t(\phi, \alpha)$ ならば、2つの平均値の差は有意である($\alpha\%$ の危険率で、平均値に差がある)と、逆に $|t_0| < t(\phi, \alpha)$ のときは、有意でない($\alpha\%$ の危険率で、平均値に差はない)と言う。

ここでは、“母標準偏差が未知で、等しいという確信のないとき”に用いられるWelchの方法⁵⁾を用い、過誤の危険率 $\alpha=1\%$ (100回の判定で1回の過誤)と $\alpha=5\%$ とした。なお、Welchの方法における t_0 値の計算式は付録-2に示す。また、 $\alpha=1\%$ では $t(\infty, 0.01)=2.58$ 、 $\alpha=5\%$ では $t(\infty, 0.05)=1.96$ である。

層別要因による平均値の差の検定に用いる2つの値は次のものとした。

- (1) 授業形式による差：講義系(全学科)と演習系(全学科)の平均値
 - (2) 学科による差：講義系と演習系のそれぞれにおいて、当該学科の平均値と全学科(当該学科を含む)の平均値
- [注]当該学科を除いた全学科の平均値と標準偏差を採るべきかもしれないが、この場合には、学科により検定対象の

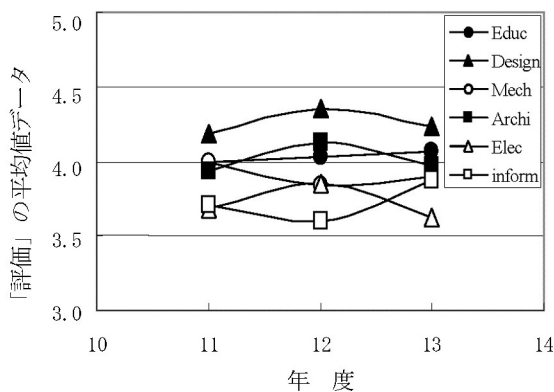


図1 「評価」の学科平均値(全授業)の時系列

全学科の値が異なることになる。全学科の値を同じにする方が良いと考えた。

3. 「評価」に対する層別因子の有意差の検定

層別因子は“授業形式(講義系授業、演習系授業)と学科”である。

3.1 授業形式による平均値の差の検定

授業形式により層別した「評価」の基本統計量(平均値 μ 、標準偏差 σ)を表2-1に、平均値の年度変化を図2に示す。同図より、演習系の「評価」は講義系の値よりも高いことが見てとれる。なお、授業形式による平均値の差は、3年間の平均値で $\mu_{DEF}=0.34$ である。

平均値の差の検定結果を表2-2に示す。同表より、授業形式による「評価」の平均値の差はほぼ確実に存在し、演習系の方が講義系より高いと言える。分散分析表は省略するが、授業形式による平均値の差が有意であることを確認している。なお、授業形式における自由度 ϕ の最小値=71.9であり、大きなデータ数に相当する。すなわち、 $t(72, 0.01) \approx 2.65$ であり、 $\phi = \infty$ での $t(\infty, 0.01)=2.58$ と同程度の検定力がある。

表2-1 「評価」の授業形式ごとの平均値と標準偏差

μ, σ, N	H11	H12	H13
平均値 μ			
講義系 (全学科)	3.74	3.83	3.80
演習系 (")	4.01	4.16	4.22
標準偏差 σ			
講義系 (")	0.51	0.50	0.50
演習系 (")	0.45	0.45	0.46
データ数 N			
講義系 (")	41	62	66
演習系 (")	35	43	44

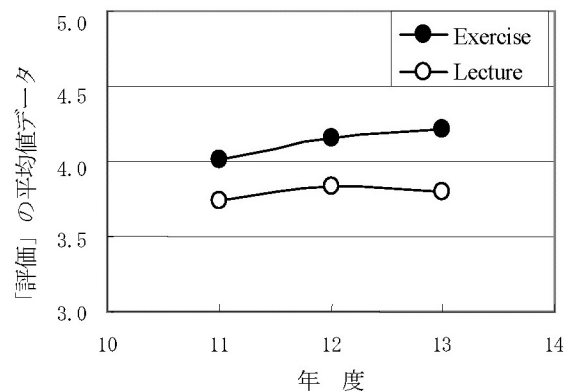


図2 「評価」の授業形式ごとの平均値(全学科)の時系列

3.2 学科による平均値の差の検定

授業形式が有意なので、講義系と演習系に層別して学科間の差を検定する。学科により層別した「評価」の基本統計量を表3-1に、平均値の年度変化を図3-1(講義系)、図3-2(演習系)に示す。

図3-1で、各学科の「評価」の年度変化は複雑に交差しており、学科間の差の判定は“図を見るだけ”では難しいことが解る。一方、図3-2では、実印と空印は明瞭に区分されており、この傾向が『数学に関係する難しい授業が多くて、学生による「評価」も悪くなりがちである』との意見になっていると思われる。しかし、学科単位で見ると、他の2つの「評価」の高い(あるいは低い)学科が全学科の平均値を上げる(あるいは下げる)ので、当該学科の平均値と全学科の平均値との差が有意になるとは言い切れないと考えられる。参考までに、2グループ間の平均値の差を検定したところ、講義系は有意ではなく、演習系のみ有意であった。

平均値の差の検定結果を表3-2に示す。同表で、 $t(\infty, 0.01) = 2.58$, $t(\infty, 0.05) = 1.96$ なので、 $t_0 < 2$ の場合は $t(\phi, \alpha)$ の記述は省略している。なお、検定のケース数は6学科×3年間=18ケースであるが、デザインのH11年度の講義系は $N = 1$ のため検定できない(標準偏差 σ が求められないため)。

- (1) 講義系：表3-2の第2欄に t_0 値を示すが、ほとんど $t_0 < 2$ である。 $t_0 > 2$ の場合を検定した結果、過誤の危険率 $\alpha = 1\%$, $\alpha = 5\%$ において、17 ケースは全て有意ではない(平均値に差はない)。

表2-2 「評価」の授業形式による平均値の差の検定結果

年度	t_0	$t(\phi, 0.01)$	検定結果
11	2.37	2.66	5%有意
12	3.43	2.63	1%有意
13	4.52	2.63	1%有意

(注)1%有意なら、当然5%有意となる。

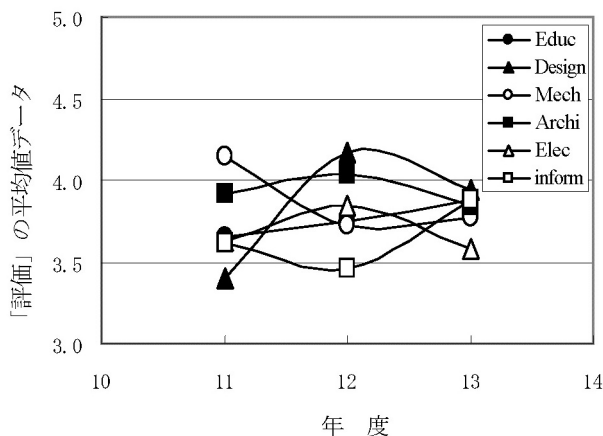


図3-1 「評価」の学科平均値(講義系)の時系列

- (2) 演習系：表3-2の第3~6欄に示すように、
- 1) $\alpha = 1\%$: 18 ケースの全てが有意ではない。
 - 2) $\alpha = 5\%$: 一般教育(H12, H13 が全学科の平均値より高い)、デザイン(H11, H12 が高い)と情報(H12 が低い)の5 ケースが有意である。一般教育の演習系は5 学科の学生が混在して行われる(学生を層別できない)ので、「評価」が高い理由は他にあるとする。一般教育を除くと、有意なのはデザインと情報の3 ケースとなる。このとき、有意なケースの比率は、演習系に限れば $3/(18-3) = 20\%$ 、全授業では $3/(17+15) = 9.4\%$ となる。

3.3 平均値の補正シミュレーション

演習系で有意となった平均値(μ_{ORG} で表す)の差は小さな偏りの場合もあることを例示する。

3.3.1 シミュレーションのモデル

優秀な学生は授業を理解できるので、概して高い「評価」を与えると考えられるし、出席も良好である。したがって、その影響はデータに含まれている。逆に、勉強困難な学生は概して低い「評価」をすると思われる、欠席する場合もある。

このように想定し、次の2つのモデルと t_0 値に関する仮定を設ける。なお、科目あたりの回答数 N_{ANC} は一般教育=20、他の学科・専攻=10 とする。

- ① モデルA：例えば、課題提出などが大変だったので、本当は「評価」=3 にしたいのだが、授業は理解できたし興味も感じたので「評価」=4 で回答したとする。1名の評点増=+1の平均値の補正値 μ_{ADJ} は；

$$\mu_{ADJ} = (\mu_{ORG} \times N_{ANC} - 1) / N_{ANC} = \mu_{ORG} - (1/N_{ANC})$$

- ② モデルB：「評価」=2~3(評点の平均で2.5)を与える学生1名が、たまたま欠席した日にアンケートをとったとする。1名の評点=2.5の補正値 μ_{ADJ} は；

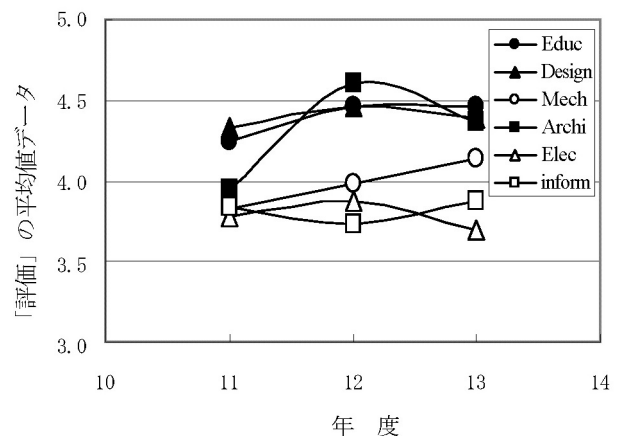


図3-2 「評価」の学科平均値(演習系)の時系列

$$\mu_{\text{ADJ}} = (\mu_{\text{ORG}} \times N_{\text{ANC}} - 2.5) / (N_{\text{ANC}} - 1)$$

③ <仮定>両モデルにおいて、 t_0 値の変化を分子(平均値の差)の変動のみと近似する(付録-2 参照)。

3.3.2 H12年度のシミュレーション

H12年度の平均値は、全学科=4.16、一般教育=4.46、デザイン=4.46、情報=3.73であり、全学平均値との差は+0.30、+0.30、-0.43である。

(1) 一般教育(平均値の差=+0.30, $N=8$)

$N=8$ の各科目において、モデルAを想定する。

各科目は $N_{\text{ANC}}=20$ なので、その平均値は+0.05があったことになる。それを補正すると、学科の $\mu_{\text{ADJ}}=4.46-0.05=4.41$ 、全学平均値との差は $4.41-4.16=0.25$ となる。仮定より、 $t_0=2.73(0.25/0.30)=2.28 > t(\phi, 0.05)=2.10$ であり、5%有意は変わらない。これより、1名の評点+1の影響を受けないこと、次いで以下に示す2学科の結果

表 3-1 「評価」の学科ごとの平均値と標準偏差

講義系	H11	H12	H13	演習系	H11	H12	H13
平均値 μ				平均値 μ			
一般教育	3.65	3.75	3.88	一般教育	4.25	4.46	4.46
デザイン	3.40	4.17	3.94	デザイン	4.33	4.46	4.39
機械	4.15	3.72	3.77	機械	3.83	3.98	4.14
建築	3.92	4.04	3.84	建築	3.95	4.61	4.37
電子	3.63	3.84	3.58	電子	3.78	3.88	3.70
情報	3.62	3.46	3.88	情報	3.84	3.73	3.87
標準偏差 σ				標準偏差 σ			
一般教育	0.230	0.441	0.455	一般教育	0.412	0.248	0.155
デザイン	----	0.534	0.336	デザイン	0.259	0.290	0.317
機械	0.403	0.283	0.401	機械	0.445	0.312	0.380
建築	0.610	0.564	0.712	建築	0.550	0.236	0.331
電子	0.494	0.391	0.580	電子	0.438	0.467	0.733
情報	0.557	0.617	0.352	情報	0.453	0.391	0.270
データ数 N				データ数 N			
一般教育	6	12	13	一般教育	8	8	6
デザイン	1	7	9	デザイン	5	11	17
機械	5	9	13	機械	5	8	7
建築	8	13	12	建築	4	2	4
電子	10	13	13	電子	5	5	6
情報	11	8	6	情報	8	9	4

表 3-2 「評価」の学科ごとの t_0 値と演習系の検定結果

学科	年度	t_0 (講義系)	t_0 (演習系)	$t(\phi, 0.01)$	$t(\phi, 0.05)$	演習 検定
一般教育	11	-0.75	1.49	-	-	
	12	-0.62	2.73	2.88	2.10	5%有意
	13	0.57	2.64	2.85	2.09	5%有意
デザイン	11	-	2.33	3.25	2.31	5%有意
	12	1.58	2.74	2.80	2.06	5%有意
	13	1.13	1.67	-	-	
機械	11	2.04	-0.81	-	-	
	12	-1.01	-1.40	-	-	
	13	-0.25	-0.47	-	-	
建築	11	0.77	-0.21	-	-	
	12	1.21	2.52	9.93	4.30	
	13	0.17	0.87	-	-	
電子	11	-0.64	-1.07	-	-	
	12	0.05	-1.25	-	-	
	13	-1.27	-1.68	-	-	
情報	11	-0.69	-0.96	-	-	
	12	-1.63	-2.92	3.01	2.16	5%有意
	13	0.50	-2.31	4.03	2.57	

(有意から有意差なしへ)を考慮すると、一般教育の「評価」は高いと考えて良い。

- (2) デザイン(平均値の差=+0.30, $N=11$)
 $N=11$ の各科目において、モデルAを想定する。
 各科目の平均値は+0.1があったことになる。補正值である $\mu_{ADJ}=4.46-0.1=4.36$, 平均値の差は $4.36-4.16=0.20$ となる。 t_0 値の変化を同様にすると、 t_0 値は $t_0=2.74 \times (0.20/0.30)=1.83$ となり、有意ではなくなる。
- (3) 情報(平均値の差=-0.43, $N=9$)
 $N=9$ の各科目において、モデルBを想定する。
 $\mu_{ADJ}=(3.73 \times 10 - 2.5)/9=3.87$, 平均値の差は $3.87-4.16=-0.29$ となる。 $t_0=-2.92(0.29/0.43)=-1.97$ となり、やはり有意ではなくなる。

3.3.3 シミュレーションのまとめ

デザイン, 情報のように、科目あたりの回答数が小さな($N_{ANC} \approx 10$)場合、1つの回答の影響は大きいことが解る。すなわち、有意となった平均値の差は1つの回答の影響による“小さな偏り”程度とも言える。

シミュレーション例より、有意な比率=20%ではあるが、回答数が小さな($N_{ANC} \approx 10$)場合を考慮して、演習系における「評価」の平均値の学科間の差は、一般教育は高く、他の学科・専攻には無いと判定する。

3.4 層別因子の検定結果

「評価」の平均値の差の検定結果より、次の知見を得た。

- (1) 授業形式による差があり、演習系の平均値は講義系の値より評点で0.34(8%)ほど高い。
- (2) 学科による差は、講義系ではない。また、演習系では一般教育は高いが、他の5学科は差がないとして良い。なお、演習系において、一般教育を除いた場合の再検定結果を5章に示す。

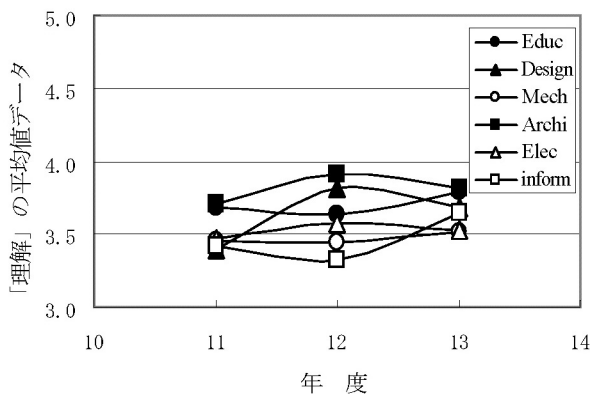


図4-1 「理解」の学科平均値(講義系)の時系列

4. 「理解」と「興味」に対する層別因子の有意差の検定

授業形式により層別した「理解」の平均値の年度変化を図4-1(講義系), 図4-2(演習系)に、「興味」の場合を図5-1(講義系), 図5-2(演習系)に示す。

講義系(図4-1, 図5-1)では、意外にも「理解」, 「興味」の平均値のばらつき幅は「評価」の幅(≈ 0.8 , 図3-1)より小さいので、有意でないことが想定される。一方、演習系(図4-2, 図5-2)では、「理解」, 「興味」の幅(≈ 1.2)は「評価」の幅(≈ 1 , 図3-2)より大きくなっている。しかし、標準偏差も大きくなっているため、 t_0 値の分子と分母が共に大きくなるので、平均値の差が大きいからといって有意とは限らないと思われる。

なお、授業形式ごとの平均値は表5-1に、学科ごとの平均値は表5-3に示す。

4.1 授業形式による平均値の差の検定

表4-1に示す検定結果より、「理解」と「興味」に対しても、授業形式は有意である。これより、講義系と演習系に層別して学科間の差を検定することになる。

表4-1 「理解」と「興味」の授業形式による平均値の差の検定結果

年度	t_0	$t(\phi, 0.01)$	検定結果
「理解」			
11	2.30	2.65	5%有意
12	3.27	2.64	1%有意
13	4.33	2.63	1%有意
「興味」			
11	3.65	2.65	1%有意
12	4.36	2.65	1%有意
13	5.00	2.63	1%有意

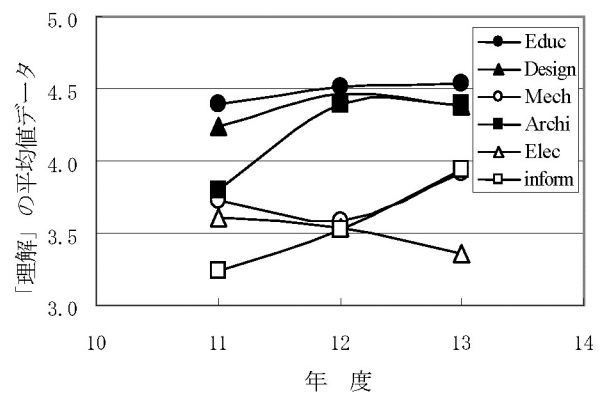


図4-2 「理解」の学科平均値(演習系)の時系列

4.2 学科による平均値の差の検定(講義系)

講義系と演習系における学科平均値の差の検定結果を表 4-2 に示す。同表の講義系の欄に見るように、 t_0 値は全て $t_0 < 2$ である。過誤の危険率 $\alpha=1\%$ 、 $\alpha=5\%$ において、17 ケースの全てが有意ではない。

4.3 学科による平均値の差の検定(演習系)

(1) 「理解」: 「評価」と同様に、一般教育を除いて有意なケースを示す。表 4-2 の演習系の欄に見るように、

- ① $\alpha=1\%$: デザイン(H12), 電子(H12), 情報(H11)の3 ケースが有意である。
- ② $\alpha=5\%$: デザイン(H11)が有意なので、合計4 ケースが有意である(1%有意なら、当然5%有意となる)。

(2) 「興味」: 同様に、

- ① $\alpha=1\%$: 有意なケースはない。
- ② $\alpha=5\%$: デザイン(H12), 建築(H12, H13)の3 ケースが有意である。
- (3) 「理解」と「興味」における層別因子の検定結果
3.3 節と同様な近似計算を行うと、これらの有意なケースは有意ではなくなると思われるので、一般教育を除く学科間の差はほとんどないと考えて良い。

5. 学科による平均値の差の再検定(演習系)

3, 4 章にて、“演習系では一般教育は高いが、他の5 学科間の差はほとんどないと良い”としたが、この前提(一般教育は高いとして除いた場合)における5 学科間の差の検定結果を示す。

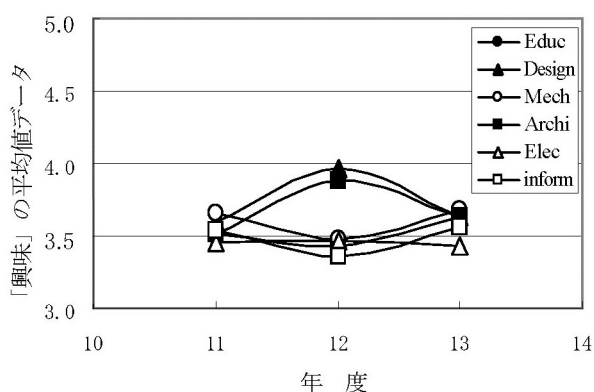


図 5-1 「興味」の学科平均値(講義系)の時系列

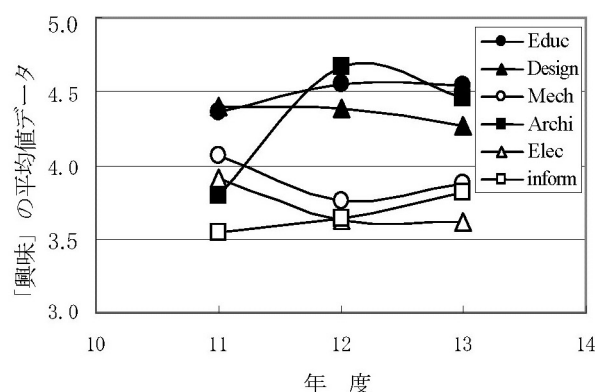


図 5-2 「興味」の学科平均値(演習系)の時系列

表 4-2 「理解」と「興味」の学科ごとの t_0 値と演習系の検定結果

学科	年度	「理解」			「興味」		
		t_0 (講義系)	t_0 (演習系)	演習 検定	t_0 (講義系)	t_0 (演習系)	演習 検定
一般教育	11	0.84	3.81	1%有意	-0.09	2.27	5%有意
	12	0.01	4.57	1%有意	-1.28	3.88	1%有意
	13	0.79	3.49	1%有意	0.18	3.61	1%有意
デザイン	11	-	2.89	5%有意	-	2.11	
	12	1.01	3.59	1%有意	1.72	2.45	5%有意
	13	0.11	1.98		0.18	1.05	
機械	11	-0.31	-0.37		0.43	0.19	
	12	-1.29	-2.21		-0.76	-1.58	
	13	-0.88	-1.36		0.68	-1.60	
建築	11	0.67	-0.12		-0.03	-0.83	
	12	1.67	1.34		1.71	4.16	5%有意
	13	0.62	1.94		0.16	2.67	5%有意
電子	11	-0.32	-1.06		-0.41	-0.80	
	12	-0.48	-3.35	1%有意	-1.25	-2.12	
	13	-0.61	-2.22		-0.77	-1.60	
情報	11	-0.49	-3.60	1%有意	0.07	-1.91	
	12	-1.31	-2.38		-0.95	-1.75	
	13	-0.06	-0.85		-0.13	-2.18	

5.1 授業形式による平均値の差の検定

「評価」、「理解」と「興味」の基本統計量を表 5-1 に、授業形式による平均値の差の検定結果を表 5-2 に示す。

表 5-2 より、

- (1) 「評価」：H11=有意差なし、H12=5%有意、H13=1%有意の時系列的な傾向も考慮すると、有意差ありとして良い。
- (2) 「理解」：「評価」と同様に有意である。
- (3) 「興味」：3 年度において、1%有意である。

以上より、「評価」、「理解」と「興味」の平均値は授業形式による差があるとして良い(3.1 節と同じ結果)。

5.2 学科による平均値の差の検定

学科ごとの平均値を表 5-3 に、平均値の差の検定結果を表 5-4 に示す。表 5-4 に示す全 45 ケースの中、次の 12 ケース(比率=12/45=27%)が有意である。

- (1) 「評価」：デザイン(H11, H12)の 2 ケース。
- (2) 「理解」：デザイン(H11~H13)、建築(H13)、電子(H12)、情報(H11)の 6 ケース。

- (3) 「興味」：デザイン(H11, H12)、建築(H12, H13)の 4 ケース。

以上より、「評価」、「理解」と「興味」の学科間の平均値の差はあるとも思える。

ここで、3.3 節と同様な平均値の補正を行ってみる。デザインと建築はモデル A、電子と情報はモデル B とした平均値の補正值 μ_{ADJ} での検定結果を表 5-5 に示す。

同表より、次のように判定する。

- (1) 「評価」：デザイン(H12)の 1 ケース(比率=1/15=7%)なので、学科間の差はないと判定する。
- (2) 「理解」：デザイン(H11, H12)の 2 ケース(比率=2/15=13%)である。デザインが他の工学科と差がないとする明確な理由を見いだすことは難しい。
- (3) 「興味」：デザイン(H12)、建築(H12, H13)の 3 ケース(比率=3/15=20%)となる。デザインは「理解」と同様である。建築は他の 3 工学科と差異があることになるが、やはり理由を見いだすことは難しい。ただ、 $N=2, 4$ と少な過ぎるので、バイアスが入りやすいとは言える。

表 5-1 授業形式ごとの平均値と標準偏差 [「評価」は表 2-1 に対応]

μ, σ, N	「評価」			「理解」			「興味」		
	H11	H12	H13	H11	H12	H13	H11	H12	H13
平均値 μ									
講義系 (全学科)	3.74	3.83	3.80	3.53	3.63	3.66	3.52	3.59	3.59
演習系 (")	4.01	4.16	4.22	3.83	4.00	4.15	4.01	4.07	4.13
(除く, 一般)	3.93	4.09	4.18	3.67	3.88	4.09	3.90	3.96	4.07
標準偏差 σ									
講義系 (")	0.51	0.50	0.50	0.57	0.52	0.61	0.62	0.50	0.59
演習系 (")	0.45	0.45	0.46	0.57	0.59	0.56	0.54	0.59	0.53
(除く, 一般)	0.45	0.46	0.48	0.53	0.59	0.58	0.55	0.59	0.54
データ数 N									
講義系 (")	41	62	66	(除く一般)					
演習系 (")	35	43	44	27	35	38	(同左)		

表 5-2 授業形式による平均値(演習系)の差の検定結果(一般教育を除く 5 学科) [表 2-2, 表 4-1 に対応]

年度	「評価」			「理解」			「興味」		
	t_0	$t(\phi, 0.01)$	検定結果	t_0	$t(\phi, 0.01)$	検定結果	t_0	$t(\phi, 0.01)$	検定結果
11	1.61	2.66		1.01			2.67	2.65	1%有意
12	2.51	2.63	5%有意	2.08	2.65	5%有意	3.12	2.65	1%有意
13	3.83	2.63	1%有意	3.57	2.63	1%有意	4.19	2.65	1%有意

参考：全学科の演習系(一般教育を含む)の場合 [表 2-2, 表 4-1 をまとめたもの]

年度	「評価」			「理解」			「興味」		
	t_0	$t(\phi, 0.01)$	検定結果	t_0	$t(\phi, 0.01)$	検定結果	t_0	$t(\phi, 0.01)$	検定結果
11	2.37	2.66	5%有意	2.30	2.65	5%有意	3.65	2.65	1%有意
12	3.43	2.63	1%有意	3.27	2.65	1%有意	4.36	2.65	1%有意
13	4.52	2.63	1%有意	4.33	2.63	1%有意	5.00	2.65	1%有意

5.3 層別因子の検定結果

5 学科(一般教育を除く)の演習系において、「評価」,「理解」と「興味」の平均値に関する差の検定結果より,次の知見を得た。

- (1) 授業形式による差がある。年度のデータ数 N で重み付けした 5 学科の演習系「評価」の平均値 $\mu_{EXE.5}$ は $\mu_{EXE.5} = (3.93 \times 27 + 4.09 \times 35 + 4.18 \times 38) / (27 + 35 + 38) = 4.08$ となるので, $\mu_{EXE.5} - \mu_{LEC} = 4.08 - 3.80 = 0.28(7\%)$ ほど, 演習系が高い。[注]全学科では, 0.34(8%)高い(3.4節参照)。
- (2) 学科による差は, 全学科の場合(3.4節, 4章)とほぼ同じ結果である。すなわち,
 - ① 「評価」: 5 学科間には, 差はないとして良い。
 - ② 「理解」: デザインは高い(良いことである)。
 - ③ 「興味」: デザインと建築は高い。建築が他の 3 工学科と異なる理由は N 数 \rightarrow 大にて検討する必要がある。

表 5-5 補正した平均値(演習系)による差の検定結果

項目	学科 年度	平均値 μ_{ADJ}	補正後 t_0	検定結果
「評価」	デザイン H11	4.23	2.06	
	デザイン H12	4.36	2.34	5%有意
「理解」	デザイン H11	4.14	3.27	1%有意
	デザイン H12	4.36	3.54	1%有意
	デザイン H13	4.28	1.55	
	建築 H13	4.29	1.54	
	電子 H12	3.64	1.64	
	情報 H11	3.32	2.08	
「興味」	デザイン H11	4.30	2.05	
	デザイン H12	4.29	2.38	5%有意
	建築 H12	4.57	4.04	5%有意
	建築 H13	4.36	2.28	5%有意

表 5-3 学科ごとの平均値 [「評価」は表 3-1 に対応]

μ	「評価」			「理解」			「興味」		
	H11	H12	H13	H11	H12	H13	H11	H12	H13
5 学科	3.93	4.09	4.18	3.67	3.88	4.09	3.90	3.96	4.07
一般教育	4.25	4.46	4.46	4.39	4.51	4.53	4.36	4.55	4.54
デザイン	4.33	4.46	4.39	4.24	4.46	4.38	4.40	4.39	4.27
機械	3.83	3.98	4.14	3.73	3.58	3.92	4.06	3.76	3.87
建築	3.95	4.61	4.37	3.80	4.39	4.39	3.79	4.67	4.46
電子	3.78	3.88	3.70	3.61	3.53	3.36	3.91	3.63	3.61
情報	3.84	3.73	3.87	3.24	3.52	3.94	3.55	3.64	3.82

表 5-4 学科平均値(演習系)の差の検定結果 [表 3-2, 表 4-2 に対応]

学科	年度	「評価」		「理解」		「興味」	
		t_0	演習 検定	t_0	演習 検定	t_0	演習 検定
一般教育	11	高いとして, 除いている					
	12						
	13						
デザイン	11	2.75	5%有意	3.96	1%有意	2.56	5%有意
	12	3.20	1%有意	4.28	1%有意	3.10	1%有意
	13	1.94		2.36	5%有意	1.48	
機械	11	-0.46		0.24		0.57	
	12	-0.83		-1.55		-0.99	
	13	-0.22		-0.98		-1.18	
建築	11	0.05		0.45		-0.43	
	12	2.85		1.72		4.70	5%有意
	13	1.06		2.31	5%有意	3.07	5%有意
電子	11	-0.71		-0.28		0.04	
	12	-0.92		-2.39	5%有意	-1.56	
	13	-1.55		-2.04		-1.39	
情報	11	-0.54		-2.55	5%有意	-1.44	
	12	-2.38		-1.75		-1.28	
	13	-2.00		-0.60		-1.68	

6. 「評価」、 「理解」のばらつきの原因に関する考察

学生と教官の双方にアンケート項目に対して個人差があるが、教官の個人差はデータのばらつきとして表れていると考えられる。H13年度のデータを例に、正規分布を仮定し、ばらつきの原因を検討してみる。

「評価」を縦軸に、「理解」を横軸にとり、両パラメータの相関関係を模した図6-1(講義系)と図6-2(演習系)を示す。図中の破線で示す枠内は、“平均値±1.5×標準偏差”内の値であり、全データ数の約87%が収まる範囲である。一方、枠外のデータは統計的には、“まれ”なので、大きく偏る原因を示唆している可能性がある。

ここでは、枠外データ(●印)の“授業計画(シラバス)”による授業内容を参照し、偏る原因に関する仮説を検討してみる。

(1) 講義系の枠外データ

1) 評点の高い方の枠外(6データ)

数学をあまり使わない科目数=3である。

<仮説1>数学をあまり使わない概論・概念に関する科目の「評価」、 「理解」は格段に良い。

2) 評点の低い方の枠外(6データ)

全て数学を多用する設計手法、専門基礎に関する科目である。

<仮説2>数学を必要とする設計手法、専門基礎に関する科目では、「評価」、 「理解」の向上は困難である。

(2) 演習系の枠外データ

1) 評点の低い方の枠外(3データ)

全て専門基礎に関する科目である。

<仮説3>専門基礎に関する科目は講義系と同様に、「評価」、 「理解」の向上は難しい。

ここでの<仮説2, 3>は文献[2]の回帰分析による仮説と一致する。また、『数学に関係する難しい授業が多くて、学生による「評価」も悪くなりがちである』との意見は、この仮説によれば、『数学を基にする“専門基礎科目”は、概論・概念に関する

科目に比べて「理解」できないため、「評価」も悪くなりがちである』と解釈すべきかもしれない。

7. 結論

授業評価のアンケート項目である「評価」、 「理解」と「興味」に対する層別要因(授業形式と学科)による平均値の差の検定により、次の知見を得た。

- (1) 「評価」、 「理解」と「興味」は授業形式(講義系、 演習系)による差があり、演習系の「評価」は5段階評点で0.34(8%)ほど高い。
- (2) 講義系の「評価」、 「理解」と「興味」は学科による差はない。
- (3) 演習系の「評価」、 「理解」と「興味」においては、一般教育の平均値は高い。他の5学科では、差のある学科はあるが、有意差なしとして良い。
- (4) 「評価」と「理解」が平均値から大きく偏ったデータから、次の仮説を導いた。
 <仮説1>数学をあまり使わない概論・概念に関する科目の「評価」、 「理解」は格段に良い。
 <仮説2>数学を基にする設計手法、専門基礎に関する科目では、「評価」、 「理解」の向上は難しい。
- (5) 数学、設計手法、専門基礎に関する「理解」はあらゆる分野において必須であり、「理解」は学生の知識獲得の目安である。「理解」向上のために、当事者である学生の自助努力・向学心を期待する。

付録-1 学生による授業評価のアンケート項目

本報においては、例えばカッコ付き「評価」はアンケート項目の評価を表す。

「評価」：この授業を総合的に見て、あなたはよい授業だったと思いますか (*=思う~思わないの5段階)

「理解」：授業の内容が十分理解できた (*)

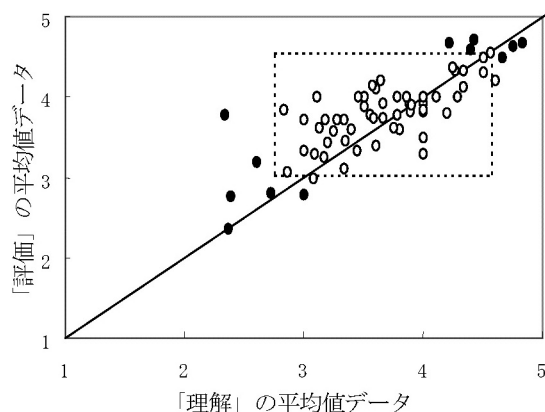


図6-1 「評価」と「理解」のデータのばらつき (H13の講義系)

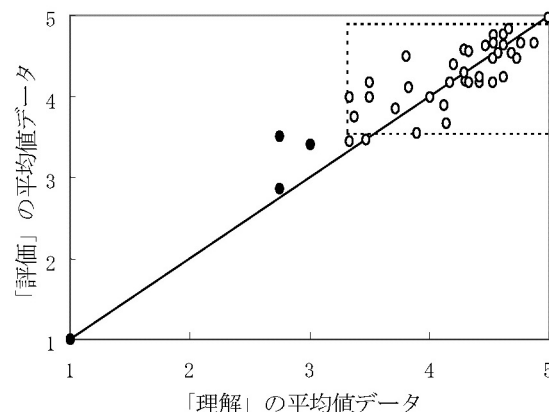


図6-2 「評価」と「理解」のデータのばらつき (H13の演習系)

- 「興味」：授業の内容があなたにとって興味あるもの感じられた（＊）
- 「態度」：この授業であなたの受講態度はどうでしたか（積極的～消極的の5段階）
- 「疎通」：先生と学生の間で、双方向のコミュニケーションが成立していた（＊）
- 「受話」：あなたは先生のコミュニケーション方法で、話を読みとれ(聞き取り)たと思いますか（＊）
- 「通話」：先生はあなたのコミュニケーション方法で、話を読みとれ(聞き取り)たと思いますか（＊）
- 「目的」：授業の目的と内容がよく一致していた（＊）
- 「機器」：黒板、ビデオ、OHPなどの情報補償機器の使用方法が適切であった（＊）
- 「準備」：テキスト・教材・実験機器などがよく準備されていた（＊）
- 「熱意」：先生の熱意を感じた（＊）
- 「程度」：授業内容の程度はあなたにとってどの位でしたか（高い～低いの5段階）

付録-2 Welchの方法における t_0 値の計算式⁵⁾

$$t_0 = (\mu_1 - \mu_2) / [\sigma_1^2 / N_1 + \sigma_2^2 / N_2]^{1/2}$$

$$\phi = 1 / [c^2 / (N_1 - 1) + (1 - c)^2 / (N_2 - 1)]$$

$$c = (\sigma_1^2 / N_1) / [(\sigma_1^2 / N_1) + (\sigma_2^2 / N_2)]$$

ここで、下添字(1, 2)は検定対象の2つのデータセット(例えば、1=講義系授業、2=演習系授業)を表し、 μ : 平均値、 σ : 標準偏差、 N : データ数、 ϕ : 自由度(データ数に相当)である。

参考文献

- [1] (私信) 学生による授業評価ワーキンググループ：平成 11 年度 聴覚部の教育活動についての点検評価—学生による授業評価—概要報告. 平成 12 年度第 11 回聴覚部教官会議資料： March 2001.
- [2] (私信) 学生による授業評価ワーキンググループ：平成 11 年度・平成 12 年度 聴覚部の教育活動についての点検評価—学生による授業評価—概要報告： Sept. 2001.
- [3] 小池将貴：5段階評定データの入力変換. 筑波技術短期大学テクノレポート Vol. 8(2)：pp.47-51, 2001.
- [4] 小池将貴：聴覚部学生による授業評価に関する一考察. 筑波技術短期大学テクノレポート Vol. 8(2)： pp.53-56, 2001.
- [5] 増山元三郎, 吉川英夫：統計解析手順集. 日科技連, p. 18, 1986.

Statistical Analysis for the Instructional Evaluation Questionnaire
by Students of Division for the Hearing Impaired
(1st Report, Findings derived from test comparing two means)

MIMAKI Toshitaro NEMOTO Masafumi NAMATAME Miki HAGITA Akio
KAWASHIMA Mitsuo WATANABE Takashi MINAGAWA Hiroki
MATSUFUJI Midori ISHIHARA Yasushi and NAKASE Koichi

Abstract : Main items of the instructional evaluation questionnaire by students are evaluation, understanding and interest. The three items are statistically tested comparing two means. Following findings are obtained.

- (1) Evaluation, understanding and interest of student are statistically significant between lecture and exercise.
- (2) For lecture, the three items have no significant difference among 6 departments.
- (3) For exercise, the three items of department of general education are higher grade than that of others, and there is no significant difference among others.

Key Words : Instructional evaluation questionnaire, test comparing two means, variance analysis