

「データの分析」の指導

— 相関を中心に —

広島大学附属中・高等学校

橋本 三嗣



高校数学教師の悩み(高1)

- 新教育課程(高等学校)

数学 I (3単位)

「数と式」、「2次関数」、「図形と計量」、「データの分析」

数学A(2単位)

「場合の数と確率」、「整数の性質」、「図形の性質」

→「データの分析」はどのように扱うべきか。

→数学Aは一年間で指導できるのか。

授業の概要



高1(数学 I)2010年10~11月

1. データの散らばり …6時間

第1時 資料(図表)の読み取り

第2時 度数分布表とヒストグラム

第3時 ヒストグラムの形状と意味

第4時 代表値(平均値, 中央値, 最頻値)

第5時 範囲と四分位範囲、箱ひげ図の形状と意味

第6時 分散と標準偏差

2. データの相関 …5時間

・第7時 散布図と相関係数

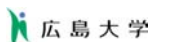
・第8時 データの収集・分析(アンケート調査)

・第9時 アンケート調査結果の発表①

・第10時 アンケート調査結果の発表②

・第11時 アンケート調査結果の発表・総括

大切にしたこと



- 生徒にとって自然な課題設定

「何か難しい話が始まった…」では残念

生徒が「やってみよう」と感じる課題を

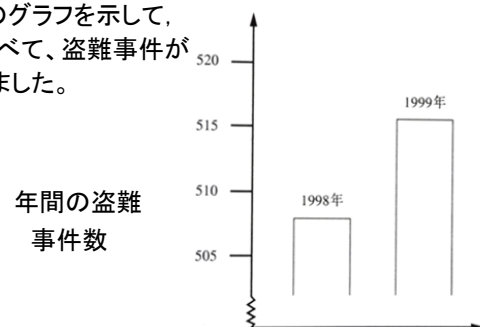
- どんな活動をさせるか

センスのない発問による活動は時間のムダ?

第1時 資料(図表)の読み取り

〔課題1〕

あるTVレポーターがこのグラフを示して、「1999年は1998年に比べて、盗難事件が激増しています」といいました。



このレポーターの発言は、このグラフの説明として適切ですか。適切である、または適切でない理由を説明しなさい。

第2時 度数分布表とヒストグラム

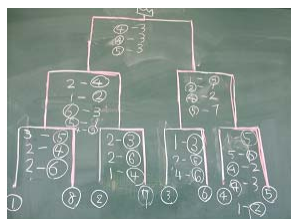


- 棒グラフとヒストグラムの違い
- 平均値の求め方

第4時 代表値

- 平均値
- 中央値(メジアン)
- 最頻値(モード)

→グラフ電卓使用

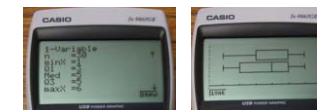


サイコロゲーム結果

第5時 範囲と四分位範囲

- 散らばりの指標(分布の幅)として
5数要約・・・データを大きさの順に並べて、その両端の最大値、最小値とデータをほぼ4分割する3つの数で構成される。この3つの数を四分位数といい、値の小さい方から、第1四分位数、第2四分位数、第3四分位数という。第2四分位数は、データを大きさの順に並べて半分に分ける中央値のことである。

→グラフ電卓使用

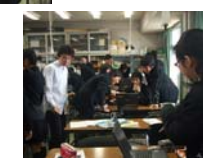
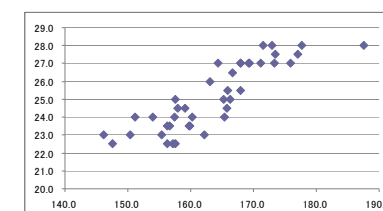


第6時 分散と標準偏差

- 散らばりの指標として
→データのすべての値を使って、データの散らばりの大きさを表す方法として
- 分散
- 標準偏差
→単位を無くし(無名数化、標準化)、解釈・比較を意味あるものにする

第8時 データの収集・分析

	x:身長cm	y:靴のサイズcm	xの偏差	yの偏差	xの偏差の平方	yの偏差の平方	xとyの偏差積
1	183.1	25.0	-0.5	0.9	0.27243	0.72873	-0.44557
2	189.4	27.0	5.8	1.9	33.3856	3.43605	10.71053
3	185.2	25.0	1.6	-0.1	2.49028	0.02141	-0.23033
4	187.5	26.0	3.9	2.9	15.3025	8.14337	48.1388
5	177.7	28.0	-14.1	2.9	198.1915	8.14337	40.1784
6	173.4	27.5	-9.8	2.4	95.61024	5.53970	23.01419
7	177.1	27.5	-3.5	2.4	18.1578	5.53970	31.72272
8	189.2	27.0	5.9	1.9	31.1483	3.43605	10.3388
9	173.3	27.0	-9.7	1.9	93.6483	3.43605	17.8388
10	185.9	25.5	2.3	0.4	5.189506	0.125074	0.805651
11	186.3	25.0	2.7	-0.1	7.171945	0.021416	-0.39191
12	184.4	27.0	0.8	1.9	0.6038	3.43605	1.442237
13	171.5	28.0	-7.9	2.9	62.9389	8.14337	22.48126
14	188.0	27.0	4.4	1.9	19.16731	3.43605	8.115407
15	188.0	27.0	4.4	1.9	19.16731	3.43605	8.115407
16	187.8	25.5	4.3	0.4	18.307	0.125074	1.512988
17	171.7	27.0	-7.5	1.9	55.921	3.43605	13.3817
18	186.7	26.5	3.1	1.4	9.474384	1.832391	4.166627
19	175.9	27.0	-12.3	1.9	150.7505	3.43605	22.75931
20	173.0	28.0	-9.4	2.9	87.9478	8.14337	26.76175
21	187.2	22.5	-6.4	-2.6	41.2469	7.003123	16.8468
22	157.4	24.0	-6.2	-1.1	38.71288	1.314099	7.132481
23	156.3	23.5	-7.3	-1.6	53.61097	2.710444	12.05449
24	157.6	25.0	-6.0	-0.1	36.2639	0.021416	0.881261
25	151.2	24.0	-12.4	-1.1	154.3089	1.314099	14.2388
26	154.0	24.0	-9.6	-1.1	92.8195	1.314099	11.03004
27	150.4	23.0	-13.2	-2.1	174.82	4.606782	28.37882
28	182.2	23.0	-1.4	-2.1	2.021945	4.606782	3.081993
29	156.3	22.5	-7.3	-2.6	53.61097	7.003123	19.3768
30	165.8	24.5	2.2	-0.6	4.743896	0.417757	-1.40776
31	180.2	24.0	-3.4	-1.1	11.70975	1.314099	3.922725
32	146.1	23.0	-17.5	-2.1	307.0188	4.606782	37.60809
33	157.9	24.5	-5.7	-0.6	32.4479	0.417757	3.68334
34	156.7	23.5	-6.9	-1.6	47.81341	2.710444	11.3959
35	157.6	22.5	-6.0	-2.6	36.2639	7.003123	15.83614
36	159.1	24.5	-4.5	-0.6	20.44804	0.417757	2.922725
37	159.9	23.5	-3.7	-1.6	13.82392	2.710444	6.17263
38	147.5	22.5	-16.1	-2.6	259.9179	7.003123	42.66419
39	185.4	24.0	1.8	-1.1	3.161457	1.314099	-2.03825
40	159.7	23.5	-3.9	-1.6	15.3517	2.710444	8.458971
41	155.4	23.0	-8.2	-2.1	67.8089	4.606782	17.84711
平均	183.6	25.1	0.0	0.0	75.9	3.4	13.9



の標準偏差 8.709722972
の標準偏差 1.842069033
r 0.86510782

課題

「I年4組生徒の生活実態」について、アンケート調査で収集したデータの中から、相関があるものをみつけ、なぜそのような関係があるのか考えよう。

個人の経験に基づく知識や知恵を多面的に用いて、データを収集し、目的に沿う適切で有効な情報を選択して利用・活用する。

発表する際の留意点

- ①作成した表やグラフなどを表示すること
- ②説明には基本的な統計量を入れること
- ③説明には自分たちの主張も入れること
- ④これまでの作業の感想も入れてもよい

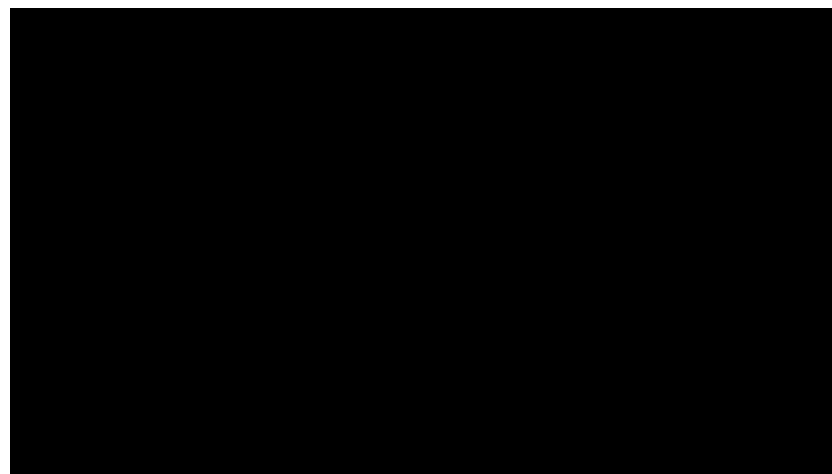
- $|r| \leq 0.2$ ほとんど相関がない
- $0.2 < |r| \leq 0.4$ やや相関がある
- $0.4 < |r| \leq 0.7$ かなり相関がある
- $0.7 < |r|$ 高い相関がある

景山三平(2007),「第6章 記述と推測の話」,『教員のための数学Ⅱ－解析・統計・コンピューター』,培風館, pp.80 - 113.

- 1班 「手の中指の長さとの長さの関係」
- 2班 「握力と体重の関係」
- 3班 「一か月に遊ぶ回数と一日にインターネットをする時間の関係」
- 4班 「一年間に読む本と漫画の関係」
- 5班 「シャトルランと中指の長さの関係」
- 6班 「運動時間と身長との関係」
- 7班 「50m走のタイムと身長との関係」
- 8班 「サンタがいないと気付いた年齢と一人で寝られるようになった時期」
- 9班 「自分の身長と相手の理想の身長との関係」
- 10班 「身長と髪の長さとの関係」



体力測定 of 相関係数



基本的な統計量・相関係数

平均 $X(\text{握力kg})=33.5$

$Y(\text{体重kg})=55.3$

男子平均 $X=41.66667$ $Y=59.8$

女子平均 $X=25.33333$ $Y=50.8$

分散 $S_x^2=99.5$ $S_y^2=102.2$

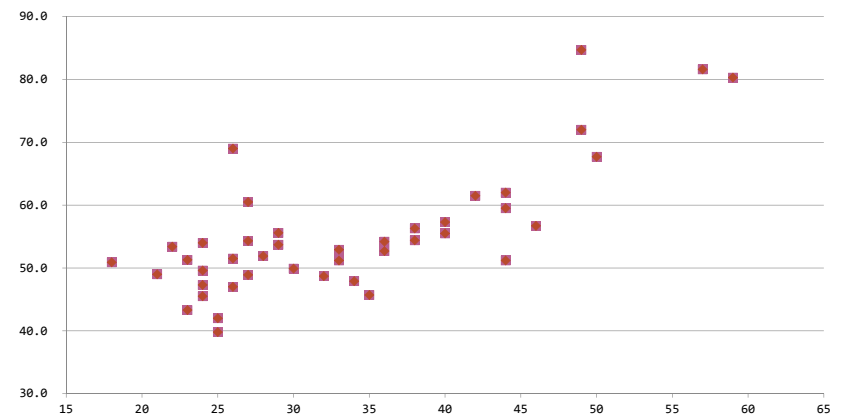
標準偏差 $S_x=9.97$ $S_y=10.11$

共分散 $S_{xy}=75.7$

相関係数 $r=0.75$

散布図

男女の散布図



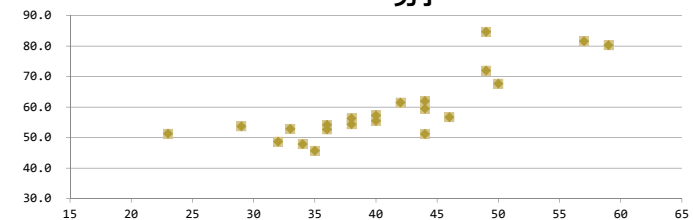
男女別の相関係数

男子	握力kg	体重kg
握力kg	1	
体重kg	0.862053	1

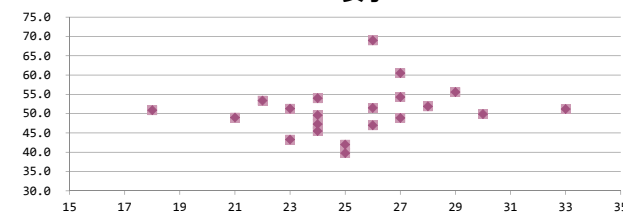
女子	握力kg	体重kg
握力kg	1	
体重kg	0.191007	1

散布図 (男女別)

男子



女子



- **体重**が大きくなればなるほど**握力**は大きくなる。**体重**が大きければ基本的に脂肪より筋肉の方が重いから筋力のある**体重**は多くなるだろう。
- 男女別によると**男子**は正の相関関係があり、**女子**は相関関係がなかった。これは**男性ホルモン**と**女性ホルモン**の特徴の違いから生まれるものである。

第10時 課題の発展

男女 (n=42)

男女	体重kg	握力kg	腕周りcm
体重kg	1		
握力kg	0.751063	1	
腕周りcm	0.831725	0.826552	1

利き腕の周りの長さ (cm)は握力(kg)と相関があるのでは？

男 (n=21)

男子	体重kg	握力kg	腕周りcm
体重kg	1		
握力kg	0.862053	1	
腕周りcm	0.824192	0.845202	1

女 (n=21)

女子	体重kg	握力kg	腕周りcm
体重kg	1		
握力kg	0.191007	1	
腕周りcm	0.760223	0.056323	1

授業を終えて(観察者からの指摘)

- 生徒発表をどのようにまとめ、データの外れ値、相関と因果関係などに関してどこまで扱うべきか。
- 生活実態の調査から相関があるものを見つけるよりも、日頃の問題意識から相関がありそうなものの中からデータを収集する方が自然である。
- 連続量と連続量のときの相関と、連続量と離散量のときの相関では意味が異なることを教師がコメントすべきであった。
- プレゼンの際には、データの数を明記するように指導すべきであった。
- 高校1年生で相関を扱う際に、グラフ(散布図)と数値(相関係数)の両面から相関を判断することが重要である。
- 教師が知らない事柄に出会った場合には生徒と一緒に考えるという姿勢が必要である。

授業を終えて(生徒の感想)

- なんか、数学の授業じゃない感じがした。
- 1回の授業で「これができるようになった」という達成感がもちにくかった。問題をたくさん解きたかった。
- 答えがはっきり決まらないものがあり、すっきりしなかった。
- 「間違えたことをいったらいけない」というプレッシャーがない授業だった。
- アンケート調査をするには、予備調査をしておくことも大切だと感じた。

- 教師が「数学的に～」とか言っているようではまだまだ未熟である
- 数学が苦手な生徒だけでなく、数学が得意な生徒に対しても「問い」次第で深まる
- 不確定的な現象を取り上げた時に、帰納的推論の中に演繹的論理をいかにして導入していくのか。科学的な結論を導く過程を学べる
- 教材を素材と扱い方の二面から再検討し、よいものを集める

ご清聴ありがとうございました