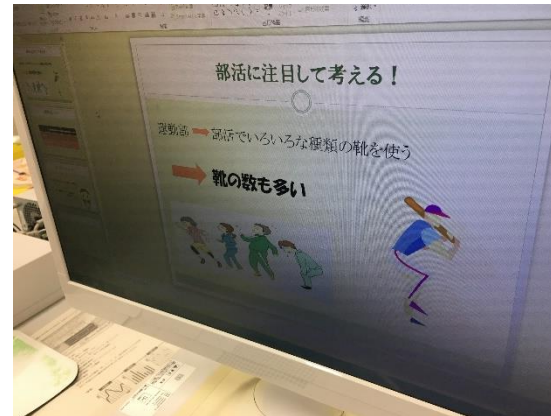




高校生が楽しく学べる 仮説検定と相関・因果

-シミュレーションから始まる“情報科”統計教育-



日出学園中学校・高等学校
武善紀之

n_takeyoshi@hinode.ed.jp

自己紹介

日出学園公式キャラクター
「日和かっぱ (ひよかっぱ)」



● 武善 紀之 Takeyoshi Noriyuki

＞ 所有免許：情報・数学・公民

＞ 教員 6 年目（**現在は情報科専任**）

日出学園中学校・高等学校（私立）



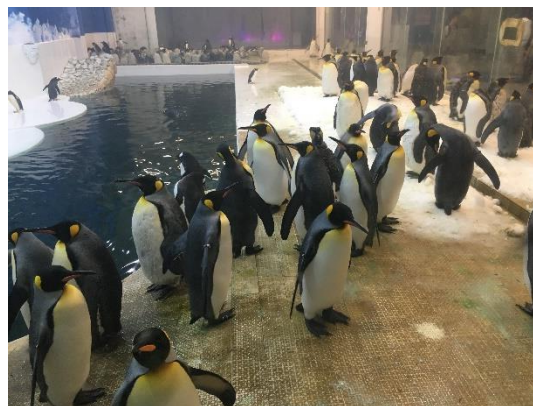
同一敷地内

- 幼稚園 101名
- 小学校 636名
- 中学高校 692名



3日前から和歌山にいます。

第12回全国高等学校
情報教育研究大会
(和歌山大会)



Agenda.

第12回全国高等学校
情報教育研究大会
(和歌山大会)



1. 実践の背景ときっかけ

2. 実践1(前). 統計的仮説検定実習(SBI)

3. 実践1(後). 統計的仮説検定実習

4. 実践2. 相関? 因果? グループワーク

5. 振り返り

Agenda.

第12回全国高等学校
情報教育研究大会
(和歌山大会)



1.実践の背景ときっかけ

2.実践1(前). 統計的仮説検定実習(SBI)

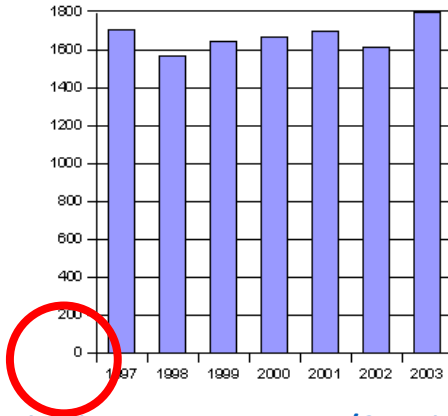
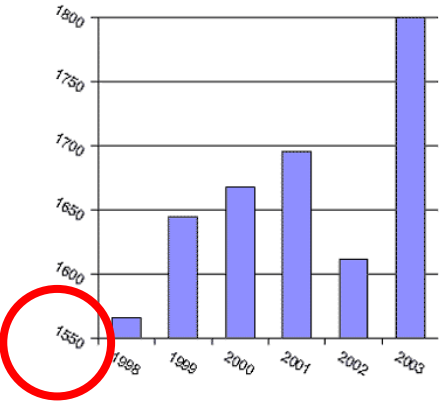
3.実践1(後). 統計的仮説検定実習

4.実践2. 相関? 因果? グループワーク

5.振り返り

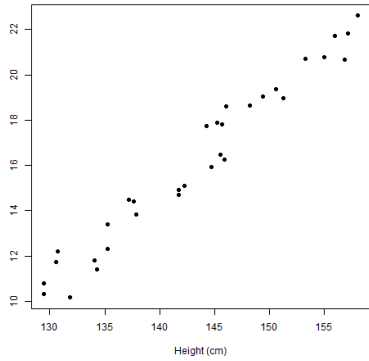
今まで

情報 ≠ PCの時間の説明に最適!



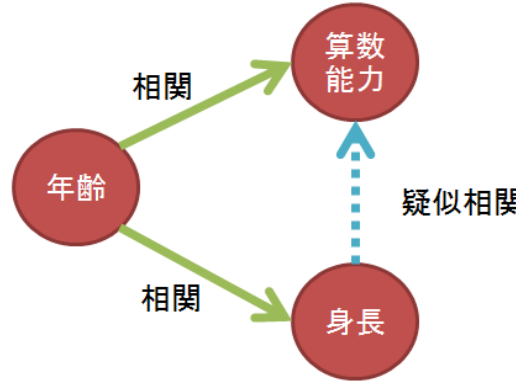
<http://psychology.jugem.cc/?eid=48>

算数の成績



身長

<http://hoxo-m.hatenablog.com/entry/20130711/p1>

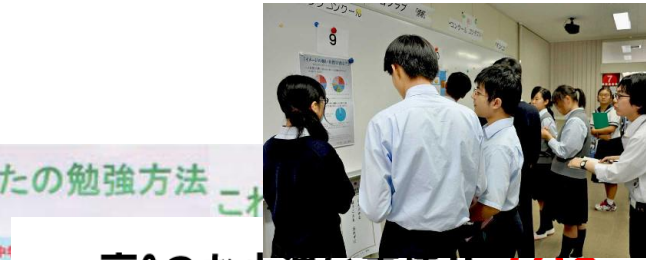


あなたの勉強方法

Q1. 日出学園 29人に

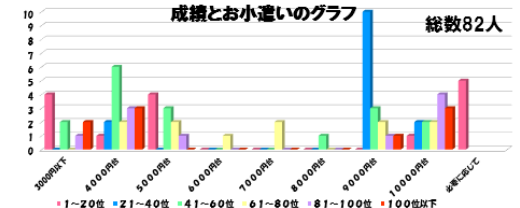
Q2. <1979年>

本番に行きたくない... 本番に行きたくない... 本番に行きたくない...



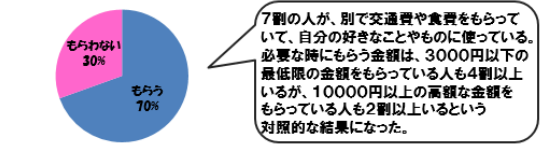
高2のお小遣い事情!! ~~100~~

大問1: 成績がいい人はお小遣いは多いのか?



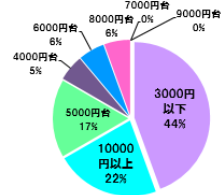
このグラフを見ると3000円以下は1~20位の成績上位者が最も多く、10000円台になると81位~成績下位者が多い傾向になっているのがわかる。

大問2: お小遣いとは別でもらうか



7割の人が、別で交通費や食費をもらっていて、自分の好きなことやものに使っている。必要な時にもらう金額は、3000円以下の最低限の金額をもらっている人も4割以上いるが、10000円以上の高額な金額をもらっている人も2割以上いるという対照的な結果になった。

大問3: 必要に応じてもらう人の金額



Point

予想では、成績上位者ほどもらっていると思ったが、予想と反する結果となった。親が甘やかしているという事が成績低下へと導いているのかもしれない...

「情報」
= パソコンの時間



「情報」の扱い方を学ぶ時間
(パソコンは手段, メタファー)



3つの目玉

プログラミング

情報デザイン

データサイエンス

学習指導要領の比較

「数学」と「情報」の
違いは…？



数学

具体的な問題の解決を通して、統計的探究プロセスを経験させるようにする

数学

複数の「質的データ」や「量的データ」が紐付けられた複数の種類のデータを取り扱う

情報

相関係数などの統計指標、相関関係や因果関係などのデータの関係性、調べようとするもの以外で結果に影響を与えている原因である交絡因子、データの関係性を数式の形で表す単回帰分析などについて扱う

数学

不確実な事象の起こりやすさに着目し、実験などを通して、問題の結論について判断したり、その妥当性について批判的に考察したりできるようにする

仮説検定の考え方

相関・因果の取り扱い



(数学で理解した内容をソフトで演習…?)



“パソコン教室”の二の舞は避けた

“情報科”的アプローチ (統計学と統計法)

- 大規模なシミュレーション
- 「帰納的」アプローチ
(cf.演繹)
- 数式に依らない説得力

[制約条件]

- **楽しく学べる**
- **説得力がある / 納得できる**
- **限られた時間数で!**
- **情報入試に対応できる!**



シミュレーションから始まる“情報科”統計教育

Agenda.

第12回全国高等学校
情報教育研究大会
(和歌山大会)



1. 実践の背景ときっかけ

2. 実践1(前). 統計的仮説検定実習(SBI)

3. 実践1(後). 統計的仮説検定実習

4. 実践2. 相関? 因果? グループワーク

5. 振り返り

検定を含んだ調査実習

高1 特進「社会と情報」、
高3 進学「情報の科学」で実施



運動部に入ると、
運動が好きになる？

	運動好き	運動嫌い
運動部所属	多	少
それ以外	少	多

テーマ決め&結果予測



プレゼンor
レポート発表



アンケート
集計&検定



プレゼンor
レポート作成

「仮説検定」の必要性

第12回全国高等学校
情報教育研究大会
(和歌山大会)



今までの統計グラフ作成では、
偶然性を排除できていない

「仮説検定」の利用場面

第12回全国高等学校
情報教育研究大会
(和歌山大会)



特進クラス
72点

進学クラス
70点

特進クラスより、
進学クラスの方が
優秀です！



え？ 本当？ たまたまじゃないの？

偶然 vs 必然(要因) のバトル

「仮説検定」の手順

第12回全国高等学校
情報教育研究大会
(和歌山大会)



Step1. 帰無仮説を設定。

A組とB組の点数には差がない (偶然生じたものだ)

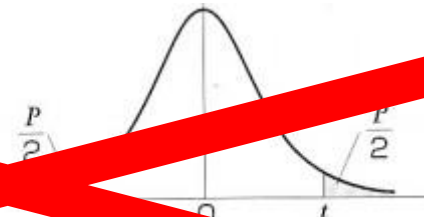
Step2. いろいろ計算する。

<http://sys-assist.co.jp/archives/1781>

偶然にしては出来すぎている、おかしいぞ

A組	B組
72	47
78	67
65	70
42	89
67	50

$$t = \frac{x_1 - x_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$
$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$



P	0.20	0.10	0.05	0.02	0.01
1	12.706	31.821	63.657		
2	6.965	17.531	31.821	24.477	31.821
3	5.841	14.267	24.477	19.162	24.477
4	5.191	12.501	20.000	16.013	20.000
5	4.773	11.578	18.505	15.085	18.505
6	4.477	10.921	17.531	14.459	17.531
7	4.257	10.521	16.993	14.067	16.993
8	4.095	10.283	16.638	13.801	16.638
9	3.982	10.133	16.449	13.641	16.449
10	3.900	10.043	16.366	13.592	16.366
11	3.851	9.988	16.312	13.556	16.312
12	3.810	9.947	16.273	13.528	16.273
13	3.776	9.914	16.240	13.504	16.240
14	3.747	9.887	16.211	13.483	16.211
15	3.722	9.864	16.186	13.465	16.186
16	3.700	9.844	16.163	13.450	16.163
17	3.681	9.826	16.142	13.438	16.142
18	3.664	9.810	16.123	13.428	16.123
19	3.649	9.796	16.106	13.420	16.106

Step3. 帰無仮説を棄却。対立仮説を採択。

仮説が間違っていたに違いない。

つまり、**A組とB組の点数には差がある[対立仮説]**



最大の障壁

理屈をどのように説明するか？



Simulation-based Inference(SBI)

シミュレーションに基づく推論

- ・ 諸外国の入門統計教育の主流
- ・ informalな統計教育

cf. formalな統計教育 (数学Bの内容的な)

確率変数や確率密度関数が必要

(出典)

Jimmy A Doi,

Active Learning Lectures for Statistical Understanding
(推測統計の概念を理解するためのアクティブラーニング授業),
第6回情報教育研究会 in 江戸川大学, July 2018

仮説検定はたったの3 Step

第12回全国高等学校
情報教育研究大会
(和歌山大会)



- 1. 仮説を立てる。**
- 2. 実験をして証拠を集める。**
- 3. 証拠に基づいて結論を出す。**

コイン投げはインチキか？



仮説 コインは公正

実験 50回このコインを投げる

予測 25回は表が出る



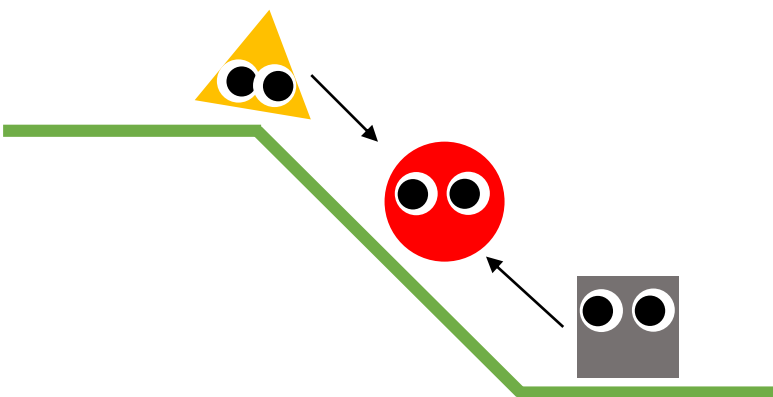
表の数	予想の範疇？	仮説は
24		
22		
2		

仮説検定はたったこれだけ！

Helper and Hinderer



- <https://www.youtube.com/watch?v=anCaGBsBOxM>
- Helper versus Hinderer
(赤ちゃんは善悪がわかるか?)



この実験を
シミュレーション
してみよう!



Helper Selected

Hinderer Selected

Total

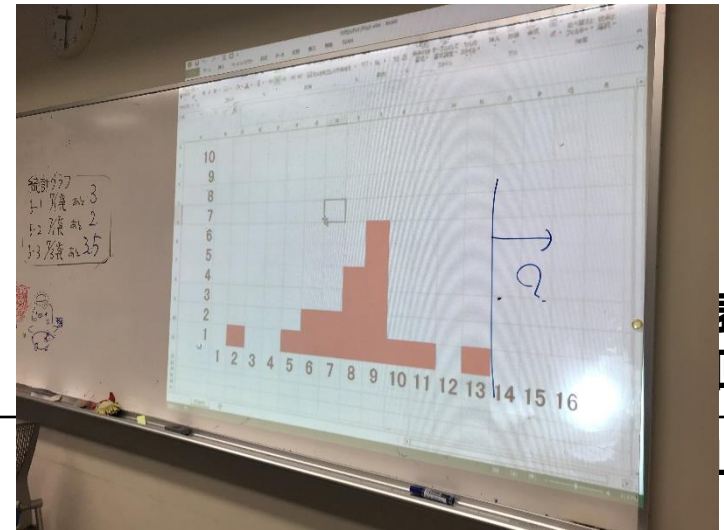
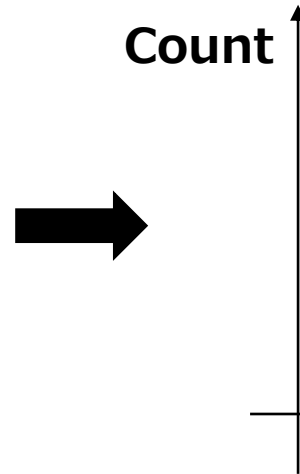
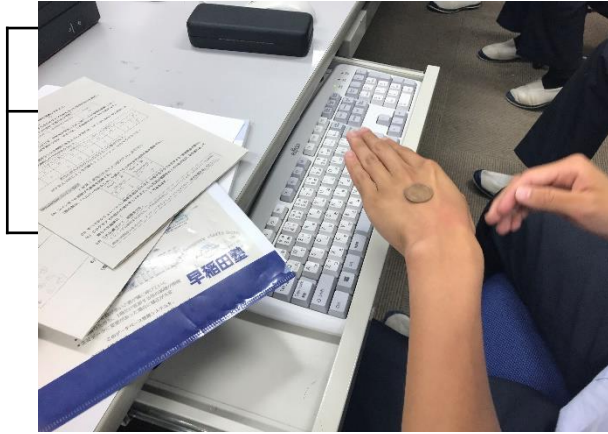
14

2

16

Helper Selected	Hinderer Selected	Total
14	2	16

Simulation 1



表の
回数
→
4

激レア（偶然ではありえないことが起こった）

14回はあり得なかった（偶然じゃない！）

➡赤ちゃんには好みがある！



Helper Selected	Hinderer Selected	Total
14	2	16

Simulation2

数学の指導要領解説ここから、
数式アプローチ（反復試行）へ！



PCを用いて、大規模シミュレーション

<http://www.rossmanchance.com/applets/OneProp/OneProp.htm>

Probability of heads:
Number of tosses:
Number of repetitions:

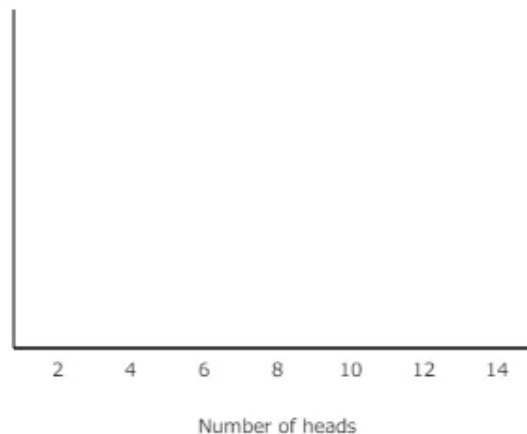
Animate

Number of heads
 Proportion of heads

As extreme as

Two-sided
 Exact Binomial
 Normal Approximation

Summary Stats



(出典)

Jimmy A Doi,
Active Learning Lectures for Statistical
Understanding

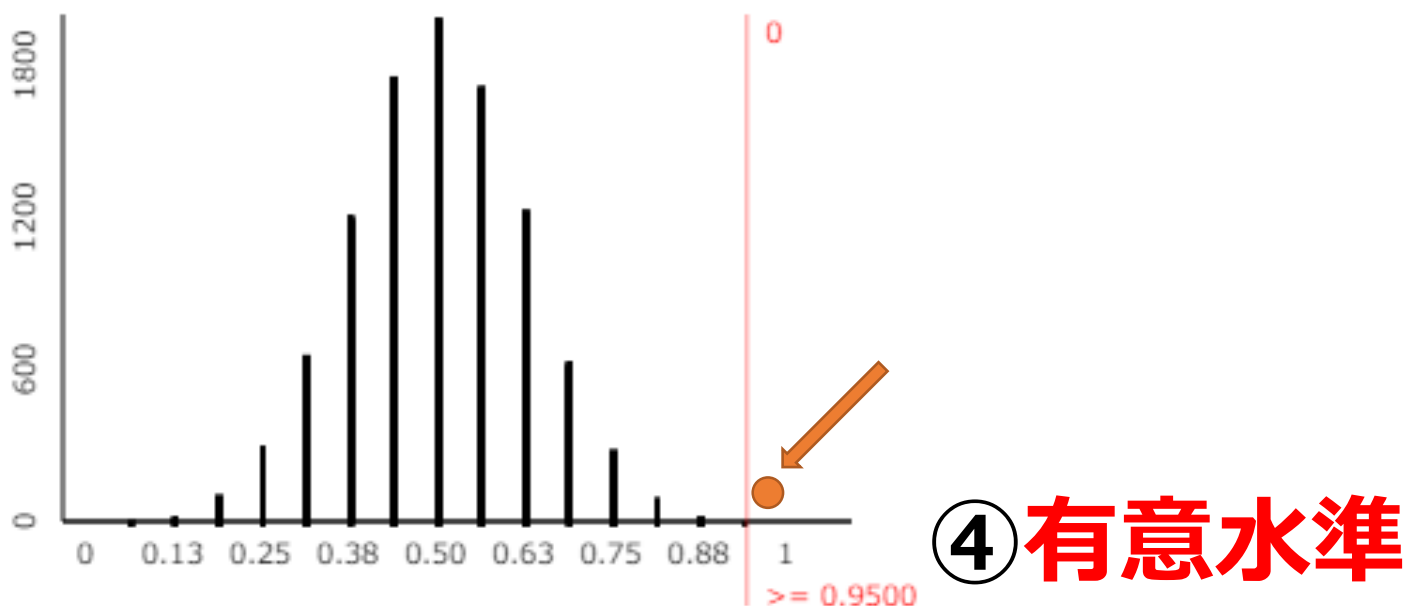
これなら全部わかる



統計的【① 仮説検定】

主張（赤ちゃんには好みがある） = ② 対立仮説

（赤ちゃんには好みがない） = ③ 帰無仮説



⑤ 棄却

⑥ 採択

Agenda.

第12回全国高等学校
情報教育研究大会
(和歌山大会)



1. 実践の背景ときっかけ

2. 実践1(前). 統計的仮説検定実習(SBI)

3. 実践1(後). 統計的仮説検定実習

4. 実践2. 相関? 因果? グループワーク

5. 振り返り



実習に使う検定は何が良い??

- 「2項検定」では、ちょっと物足りない。
(ある/なし, 勝ち/負け, ……)
- でも、こみいった内容は使いたくない。
⇒ 「何となくでもわかる感」

フィッシャーの直接確率計算

フィッシャーの直接確率計算

フィッシャーまで扱えるのは、
情報科だからこそ！



所属部活 \ 好み	好き	嫌い
運動部	9	6
運動部以外	5	10

← この表がどれくらい
期待度数表から偏ったものか
を考える

(期待度数表)

所属部活 \ 好み	好き(14)	嫌い(16)
運動部(15)	7	8
運動部以外(15)	7	8

10%以下のレア度なら、
帰無仮説を棄却！
(有意傾向も可とした)

$$\begin{array}{ccccccc} \text{クラスの人数} & \times & \text{運動部の割合} & \times & \text{好きの割合} & & \\ 30 & \times & \frac{1}{2} & \times & \frac{14}{30} & & \end{array}$$

【参考】大貫和則

統計リテラシーを育成するアンケート
調査実習の実践と課題

第4回全国高等学校情報教育研究大会
(2011)



「仮説検定」を使って、 クラスの実態を調査してみよう！

(2x2のクロス集計表に落とし込む)

疑問 運動部の人ってみんな体育の時間が好きだ

仮説 運動部に入ると、みんな好きになる？

予測

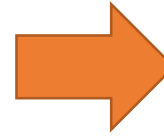
所属部活\好み	好き(14)	嫌い(16)
運動部(15)	7	8
運動部以外(15)	7	8

この表のレア度を判断すればよい

実習のプロセス



	運動好き	運動嫌い
運動部所属	①	②
それ以外	③	④



- ① 運動部所属で、運動が好き
- ② 運動部所属で、運動が嫌い
- ③ 運動部以外で、運動が好き
- ④ 運動部以外で、運動が嫌い

【特進】仮説検定実習2018_1

※切は9/18(火)です。匿名収集ですので、よろしくお願ひします。

このフォームを送信すると、メールアドレス (n_takeyoshi@hinode.ed.jp) が記録されます。自分のアカウントでない場合は、[アカウントを切り替えてください](#)

*必須

1 *

- ① 男子でおみくじの内容を信じる
- ② 男子でおみくじの内容を信じない
- ③ 女子でおみくじの内容を信じる
- ④ 女子でおみくじの内容を信じない

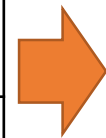
2 *

- 1 あなたは、内部進学生で、給食よりお弁当のほうが好きですか。
- 2 あなたは、内部進学生で、お弁当より給食のほうが好きですか。
- 3 あなたは、外部進学生で、給食より、お弁当のほうが好きですか。
- 4 あなたは、外部進学生で、お弁当より給食のほうが好きですか。

3 *



所属部活 ＼好み	好き	嫌い
運動部	12	6
運動部以外	5	14



年間の水族館・動物園訪問回数
年間10回以上 年間10回未満

知入

考察

- そもそも3種類が厳しい
- 好き=詳しい?
- 好きだから増えるを聞いたかったが、行くから詳しくなったの関係かも?
- その他の動物について聞いてみると、関係しない動物もある?

Rで検定

```
> mx=matrix(c(12, 6, 5, 14),
nrow=2, byrow=T)
> fisher.test(mx)
```



Fisher's Exact
p-value = 0.02171



p<.05より有意

プレゼン評価シート(4-4)

このフォームで送信すると、メールアドレス (n_takeyoshi@hinode.ed.jp) が記録されます。自分のアカウントでない場合は、[アカウントを切り替えてください](#)

*必須

評価対象者 *

選択

発表内容 *

発表者のいた形跡が認められず、中身でない	<input type="radio"/>	<input type="radio"/>
----------------------	-----------------------	-----------------------

研究の妥当性、解釈 *

妥当性がなく、結果の解釈も不適当	<input type="radio"/>	<input type="radio"/>
------------------	-----------------------	-----------------------

発表技術①発表の構成 *



Google Formを使って、
5分に短縮
匿名性も保たれる
(テストの成績等)

プリント



【データの集計】
2×2クロス集計表

属性		検定	
一軒家	マンション	条件	マンションに傾き
		① 正正正	正T ③
		②	正 ④ 正 ①

【統計ソフトRの活用方法】
Rのコマンド
\$ x <- c(1,2,3,4,5)
\$ ave(x)
\$ median(x)
\$ var(x)
シャープの正確確率検定
prop.test(c(2, 10, 15, 3), nrow=2, byrow=T)
Test for Count Data
起こりうる確率
p<0.01 有意である
0.05 有意である
<0.10

難しいところ（ある意味,醍醐味）

第12回全国高等学校
情報教育研究大会
(和歌山大会)



【質的調査の難しさ】

- ・ 運動が「好き」とは？
- ・ 成績が「良い」とは？

主観を如何にして、
客観視できる数値に落とし込めるか

- ①運動部所属で、運動が好き
- ②運動部所属で、運動が嫌い
- ③運動部以外で、運動が好き
- ④運動部以外で、運動が嫌い

- ・ 好きな教科の3番以内が体育
- ・ 模試でクラス順位が半分より上
- ・ 週3回以上自主的に運動する

【全員有意差は出ない】

11名/28名が有意

有意差が出なくても…

- 仮説が間違っていた？
- 実験計画上の問題？
 - ・ 聞き方？
 - ・ 質問内容？

「考察」できる点は、
たくさんある！

プレゼンテーション例

第12回全国高等学校
情報教育研究大会
(和歌山大会)



考察 (先に食べるタイプ→「先」、後で食べるタイプ→「後」とする)

- 好きな食べ物がシチューなどの温かいものだった場合、**冷めてしまう**ため必然的に「先」になる
→ 食べ物の種類を絞る(例・好きな野菜)
- 兄弟がいる人だけで見ると圧倒的に「先」が多い
- そもそも「後」のほうが少数派である可能性
- 兄弟がいる人よりも、いない人のほうが多すぎる
→ 同数でデータを集計すれば・・・
- 兄弟がいらないということは、**好きなものを好きなタイミングで食べることが可能**であるということ
→ 必ずしも「後」だとは限らない

なぜこのような結果に

- 目的地までの距離は？
- 車移動？電車移動？
- その時の状況によって判断が変わってくる

→ **具体的な設定が必要**

- 「好きなもののプレゼン」と違い、**絶対コピペにならない。**
- 自分の立てた仮説の為に、**生徒が必死に考察を考える**

Agenda.

第12回全国高等学校
情報教育研究大会
(和歌山大会)



1. 実践の背景ときっかけ

2. 実践1(前). 統計的仮説検定実習(SBI)

3. 実践1(後). 統計的仮説検定実習

4. 実践2. 相関? 因果? グループワーク

5. 振り返り

しかし、武善はしつこかった

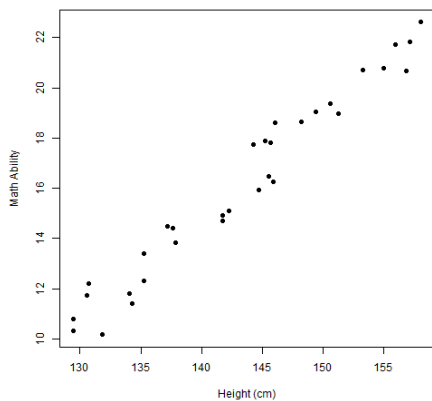


本当に「運動部に入ると、
運動が好きになる」？

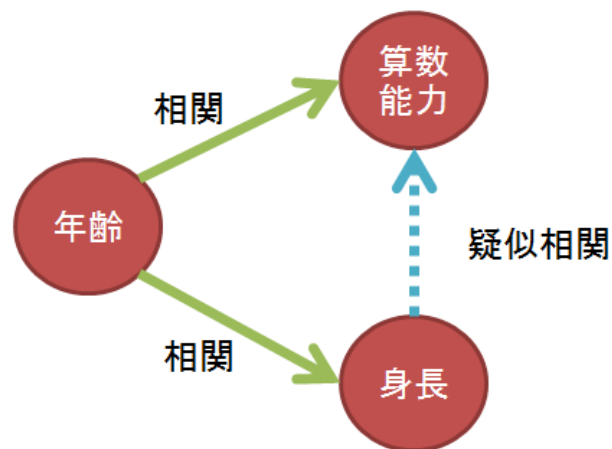
運動が好き**だから**、運動部に入ったんじゃないの？

親がスポーツに熱心**だから**、運動部に入部させて、
子供も運動を好きになったのでは？

算数の成績



身長





「検定」を扱うなら、
相関・因果の見極めはセットで
必ず扱うべきではないか？

クロス集計と相関・因果は入試頻出

第12回全国高等学校
情報教育研究大会
(和歌山大会)



大学名	問題内容
明治大学 情報総合2018	学校で実施したアンケート調査に対するクロス集計表について、 ϕ 係数を用いて分析。
明治大学 情報総合2016	健康食品広告の表記に関して、クロス集計表に生じる大小関係、サンプリングバイアス等。
慶應義塾大学 商学部2016 論文テスト	クロス集計表に対する χ^2 乗検定
慶應義塾大学 総合政策学部2017 小論文	糖尿病の死亡率と平均年収の関係について、さまざまな要因を考慮して図示
慶應義塾大学 商学部2017 論文テスト	「因果関係の成立要因」に関する文章題 (ex. 鶏が先か卵が先か?)
センター試験試作問題 (数学) 2013	「相関係数が高い場合、因果関係もある」の正誤

**「仮説検定」「相関・因果」「クロス集計」は、
入試に出しやすい。
情報入試でも今後主流に…?**

ピックアップ

この2つを組み合わせると、
ゲーム感覚にしたら面白いのでは…？



慶應義塾大学 商学部2017 論文テスト [相関と因果の違いの定義]

① 出来事の共変

出来事Xと出来事Yは一緒に（共に）変化しなければならない（相関関係）。

② 時間的順序関係

XがYより時間的に先に起こっていないなければならない。

③ 他の原因の排除 (交絡因子, 第3変数)

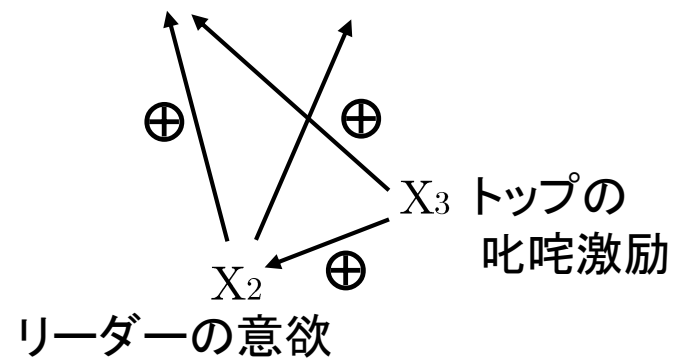
出来事X以外にYを説明できるものが何も存在しないこと。

E.B.ゼックミスタほか、
クリティカルシンキング 入門篇、
北大氏書房

慶應義塾大学 総合政策学部2017 [相関関係・因果関係の可視化]

職場にコーヒーマーカーを置くと、
生産性が向上する！

生産性 Y \longleftrightarrow X₁ コーヒー消費量



表記ルール

因果関係 A(原因)→B(結果)

相関関係 A \longleftrightarrow B

A増,B増の時は⊕

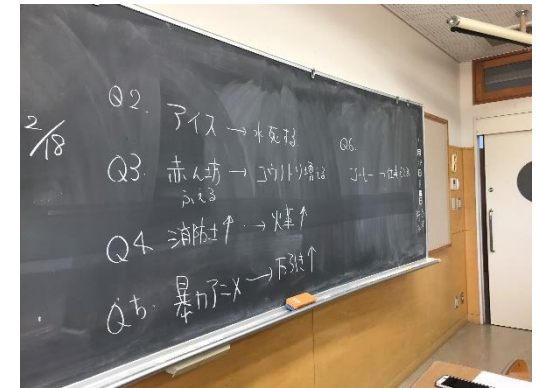
A増,B減の時は⊖



グループワークで入試問題を解く

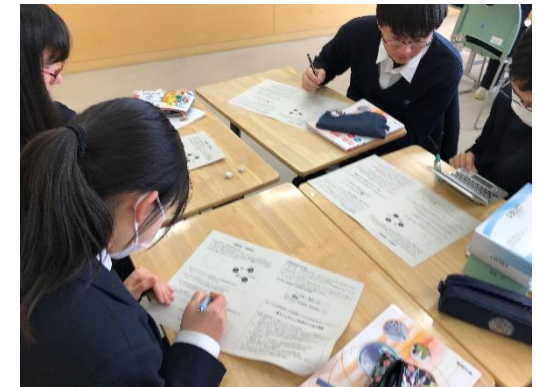
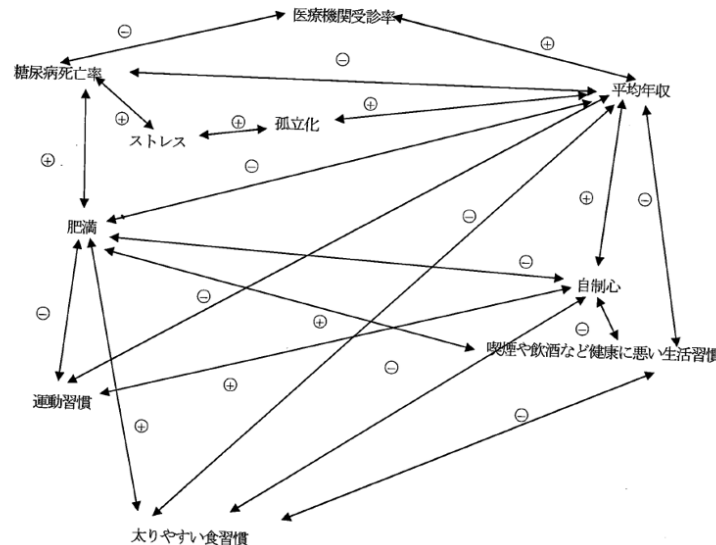
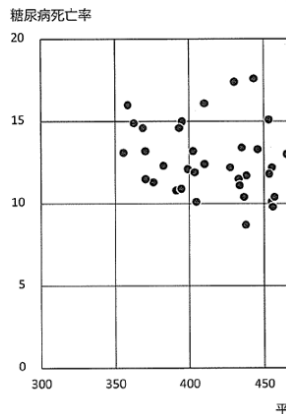
疑似相関クイズ

- Q1.小学生の身長と算数の成績
- Q2.アイスの売り上げと水死者数
- Q3. 消防士が出勤するほど、火事が大きい
- Q4.暴力シーンを見ると、非行に走りやすい
- Q5. コーヒーメーカーを置くと、生産性が向上する！



慶応大学の入試問題に挑戦！

問2 図1は都道府県の成人男性（65歳未満）の糖尿病の死亡率（人口10万人当たり死亡人数）と平均年収（百万円）を散布図にしたものです（データは仮想です）。各都道府県の年齢構成は同一となるよう調整してあります。ここでは糖尿病の死亡率が最終的な結果だとします。問1の回答および資料5～7を踏まえ、解答例1（河合塾作成）の死亡率と平均年収の間の関係の構造を図示図）→B（結果）、相関関係を示すときにはAから⊕、Aが増える時、Bは減るなら⊖をつけません。なお、図示化の例は資料3の中にあり



(出典)河合塾

自分たちの研究で！



自分たちの研究結果から因果関係が本当にあるのか考えてみよう！

一番もっともらしい提案をした班が優勝

有意差（連関があったと認められた）が出た研究結果

(1) 「男女」×「Twitter と Instagram の使用率」(研究員の調査結果)

Twitter は主に文章、Instagram は主に写真。情報収集や自分の考えを広めたい男性は Twitter に、暇潰しや様々な人と繋がりたい女性は Instagram に流れる？

(2) 「男女」×「Instagram の使用率」(研究員の調査結果)

Instagram を使うことは女子力の象徴である。Instagram を使用することで、女子力がアップする。

(3) 「妹弟の存在」×「子ども好き」(研究員の調査結果)

妹弟がいると、子どもに慣れることができるので子どもを好きになる！

(4) 「運動部／運動部以外」×「体育の好み」(研究員の調査結果)

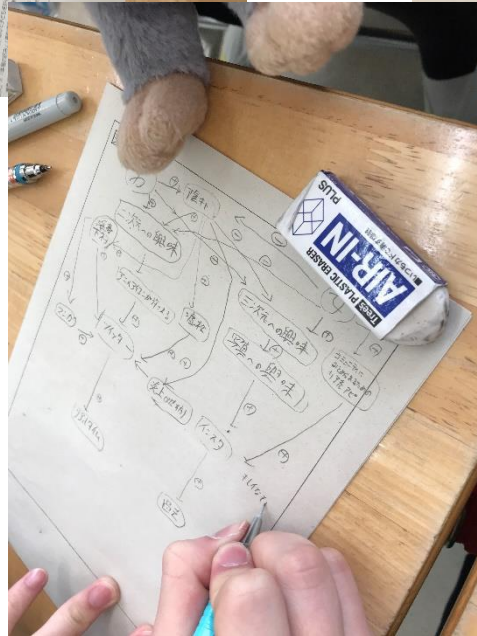
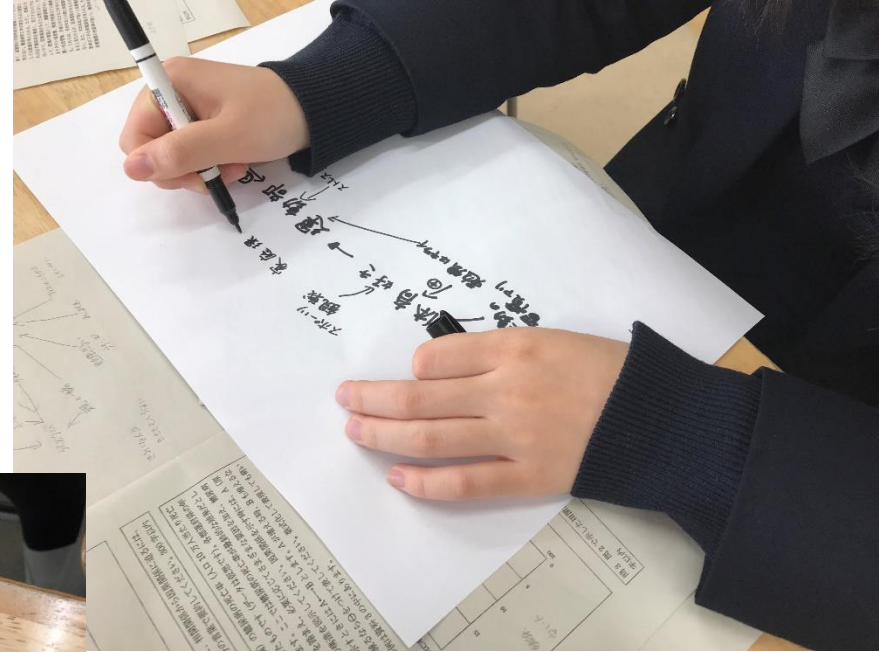
体育が好きだから、運動部に入部する！

(5) 「男女」×「メイク／すっぴんの好み」(研究員の調査結果)

男子には関係が薄く近寄りやすいイメージをメイクは生じさせるが、女子にとってはかわいく生まれ変わりたいという事実に共感できる！

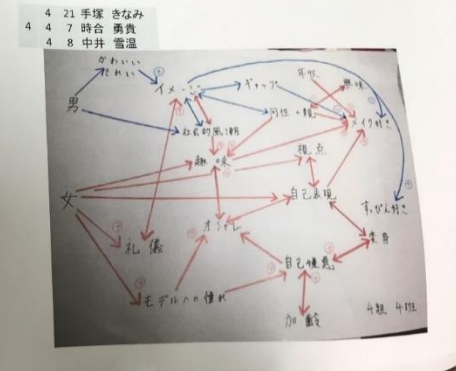
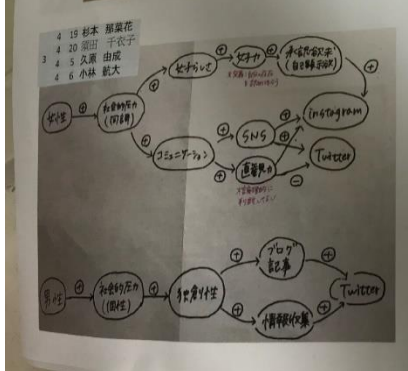
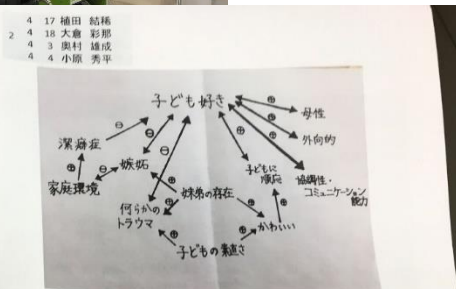
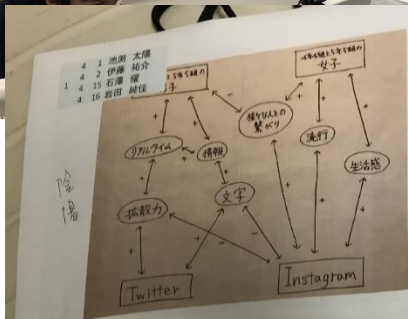
グループワーク

第12回全国高等学校
情報教育研究大会
(和歌山大会)



発表・評価

第12回全国高等学校
情報教育研究大会
(和歌山大会)



相関関係・因果関係を図示しよう ～1番説得力があるのは誰だ～

番 氏名

- 有意差(連関があったと認められた)が出た研究結果
 (1)「男女」×「TwitterとInstagramの使用率」
 (2)「男女」×「Instagramの使用率」
 (3)「姉弟の存在」×「子ども好き」
 (4)「運動部/運動部以外」×「体育の好み」
 (5)「男女」×「メイク/すっぴんの好み」

班 組 番 氏名	選択 テーマ 1～5から	内容(3点満点)				発表スキル(2点満点)			合計 (18点)	投票 (◎)
		着眼点	説得力	準備 (丁寧さ)	質疑 対応	声	視線	身振り 手振り等		
1										
2										
3										
4										
5										
6										
7										

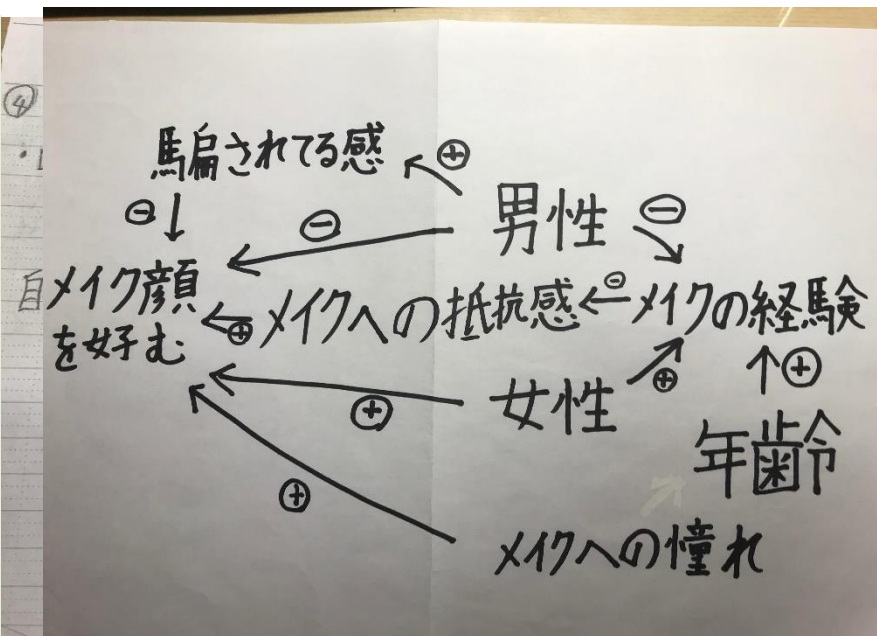
評価用 ルーブリック	3	2	1	0
着眼点	新しい要素、関連が示され、論拠の中で十分に活用されている。	新しい要素、関連は示されているが活用されていない。	当たり障りの無い要素、関連を示すだけに留まっている。	発表資料が揃っていない、発表できなかった。
説得力	納得せざるを得ない結論と図示である。	一応の根拠としては成り立つ図示と論拠である。	図示と説明がらぐで噛み合っていない、説明に説得力が無い。	発表資料が揃っていない、発表できなかった。
準備 (丁寧さ)	十分に時間がかけられ、要素の抜けモレ等もない。	3と2の間	まだまだ付け足すべき要素、示さなければいけない要素がいくつかある。	発表資料が揃っていない、発表できなかった。
質疑 対応	説得力のある回答を、質問者の理解を促す形で示すことができた。	回答はしていたが、説得力に欠ける。	回答はしていたが、明らかに的を外れて要領を得ない。	回答できなかった。
声	全ての発表者の声がよく聞かれる。	どこどころ聞こえないときがあった。	ほとんど聞こえなかった。	
視線	発表者と何度も目が合い、会場全体に視線を配っていた。	上げようという努力は感じられるが、全体としてほとんど目があらなかった。	原稿を読み続けるだけで、発表者があやうくない。	
身振り 手振り等	聴衆に理解を促そうという工夫が十分に感じられた。	努力は感じられるが、効果的ではない、あるいはほとんど示されていない。	原稿を読み上げるだけ、棒立ちのまま発表するだけ。	

ピックアップ1



男子より女子の方が、メイクされた顔を好む。

- 「校則により禁止されること」で、化粧へのイメージが低下している？
(特に調査対象が高校生である為)
- すっぴんを好む
(好むと魅力は別？)
- メイクをする場合は主観評価、
メイクをしない場合は客観評価
- 主観評価の上で、
メイク＝自身や自己評価の向上
＝「心を明るくする」行為



結論 すっぴん・メイク顔の好みは男女差はあっても
経験・年齢などの相関関係から生まれたもの。
将来変わっていく可能性大。

Agenda.

第12回全国高等学校
情報教育研究大会
(和歌山大会)



1.実践の背景ときっかけ

2.実践1(前). 統計的仮説検定実習(SBI)

3.実践1(後). 統計的仮説検定実習

4.実践2. 相関? 因果? グループワーク

5.振り返り

生徒の理解度調査(2クラス)

第12回全国高等学校
情報教育研究大会
(和歌山大会)



定期試験にはこうやって出してみました。出題例(日出国 高一特進クラス 2学期期末試験)

I. 統計的仮説検定に関する次の文章を読み、問いに答えなさい。

データを使用する多くの研究で生じる問題を、統計的仮説検定は解決してくれる。どのような調査結果も偶然で起こり得るのであれば、単なる偶然でそのような結果が得られる確率を実際に計算するのである。その確率があらかじめ定めた基準(通常, 0.05)よりも低ければ、偶然の結果とみなさず、何か意味のある(=有意な)結果が得られたと考える。例えば「店舗の改装の有無により、売上が異なる」と解釈できるデータを得たとしよう。この解釈を統計的に検証するためには、「【 A 】」と、まずは仮定する。そう仮定することで、実際に観察された売上の差が単なる偶然によって生じる確率を計算できる。この例のように、「観察された差は単なる偶然の産物であり、意味のある差ではない」と仮定する考えは帰無仮説と呼ばれる。一般的に、帰無仮説に基づいて計算された確率が 0.05 を下回る時にのみ、すなわち「単なる偶然では 20 回に 1 度あるはそれ以下の確率でしか生じない事象が起こった時」のみ、帰無仮説を捨て去る(棄却する)。棄却した結果平均以上か下回るのが、自分が本当に主張したかった【 B 】(対立仮説)である。

A 君が行ったウェブサイトのデザインを検討する為の A B テスト(ウェブサイト内の一部分を変更し、パーツ A と B のどちらがより優れているかを検討するテスト)の結果について考えてみよう。この場合の帰無仮説は「ボタンの色とボタンの押されやすさには関係が無い」になる。この帰無仮説を、表 1 のようなデータに対して検定する際、カイ 2 乗検定という手法がよく用いられる。

表 1: 元データ表(実験値 O)

	ボタンを押した	押さなかった	合計
青いボタン	70	180	250
赤いボタン	30	120	150
合計	100	300	400

カイ 2 乗検定ではまず、「もし関係が無かったら、きつこうなるだろうという回数」を求める。これを期待度数と呼ぶ。期待度数の求め方としてまず、青いボタンに注目する。青いボタンを目にした人の数は「250 人」であり、次に「ボタンの色に関わらない、ボタンを押した人の割合」を見る。すると、全体 400 人のうち、100 人がボタンを押していることがわかる。その為、「もしもボタンの色と押されやすさに関係がないのだとしたら、4 分の 1 の割合でボタンが押される」ということが言える。すなわち、「もしもボタンの色と押されやすさに関係がないのだとしたら、青いボタンを押す人の数は、 $250 \div 4 = 62.5$ 人になる」と計算できる。この 62.5 人が期待度数である。同様に計算すると期待度数表は次のようになる。

表 2: 期待度数表(期待度数 E)

	ボタンを押した	押さなかった	合計
青いボタン	62.5	187.5	250
赤いボタン	37.5	112.5	150
合計	100	300	400

その後、元データと期待度数の違いを以下の式を使って計算する。

$$\frac{(\text{元データ} - \text{期待度数})^2}{\text{期待度数}}$$

すると、次の表を得ることが出来る。

表 3: 計算結果

	ボタンを押した	押さなかった	合計
青いボタン	0.9	[F]	250
赤いボタン	[G]	[H]	150
合計	100	300	400

0.9, [F], [G], [H] の 4 つの値の和がカイ 2 乗値と呼ばれる。このカイ 2 乗値の性質はよく調べられているので、統計学の教科書を参照すれば、口値を知ることができる。実際値口が期待度数口と異なっているほど、カイ 2 乗値は【 I 】、口値は【 J 】。今回の場合、カイ 2 乗値が 7.88 を超えれば、 $\square < 0.05$ で有意、すなわち帰無仮説を捨て去ることができる。すなわち、今回の仮説検定の結論は【 K 】ということになる。

(慶應義塾商学部入試問題(H28)および <https://logics-of-blue.com/chi-squared-test/>より改題)

ト

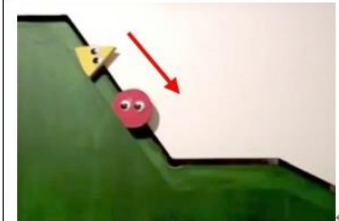
- 問1 帰無仮説【 A 】、対立仮説【 B 】をそれぞれ簡潔に述べよ。58%, 58%
(※A:店舗の改装の有無によらず、売上には差がない B:店舗の改装の有無は高利上げに関係がある。)
- 問2 【 C 】 ~ 【 H 】に当てはまる数値を入れよ。85%, 81%, 77%, 38%, 50%, 42%
(※【 C 】 187.5 【 D 】 37.5 【 E 】 112.5 【 F 】 0.9 【 G 】 1.5 【 H 】 0.5)
- 問3 【 I 】 と 【 J 】 に入る語句を次の選択肢からそれぞれ選べ。

【統計的仮説検定】

問4 1 歳未満の子供たちは、親しみやすい振る舞い (Helper) と邪魔者の振る舞い (Hinderer) の違いを認識、
問5 すなわち善悪の判断を行うことができるだろうか? (実験参加者: 16 人の 10 ヶ月の幼児)

各々の幼児は、「登山者」キャラクターを 2 試行提示される。

各々の幼児は、2 つのシナリオを何度か見せられる。
Helper によってトップへ。
Hinderer によって底の方へ。



●各々の幼児が、Helper と Hinderer のどちらを選ぶか記録を取る。



実験結果	Helper Selected	Hinderer Selected	Total
	14	2	16

【特進コース】 「クロス集計表」に対する x2乗検定(慶應の改題) (発展形)

【進学コース】 「Helper」 vs 「Hinderer」 問題のアレンジ(復習系)

正答率(参考程度)

第12回全国高等学校
情報教育研究大会
(和歌山大会)



問題の種類	新規出題	復習
帰無仮説を述べる問題	61%	43%
対立仮説を述べる問題	61%	43%
有意差ありの場合の結論を述べる問題	—	62%
有意差なしの結論を述べる問題 (“有意差がない”⇒“差がない”にしてしまう)	27%	38%
別の要因を発見する問題	—	67%
相関と因果の違いを説明する問題 (別の問題より)	85%	—

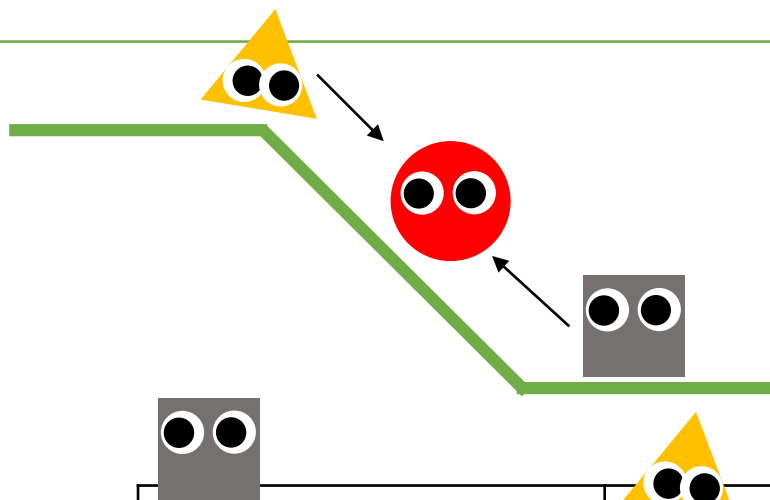
- ・ 「相関」と「因果」が異なることは明確に理解できている
- ・ 有意差があれば、差があることも説明できる
- ・ **有意差がない場合 (つまり、本質的な偶然vs必然ロジック)**
の解釈に課題

案外できていた問題



検定結果が有意であったとき、「赤ちゃんには善悪の区別がついている」とこの実験だけから本当に判断はできるだろうか？

他の要因は完全に排除できているか、他の要因について1つ挙げ、どのような実験を行えばその要因を排除できるか述べよ。



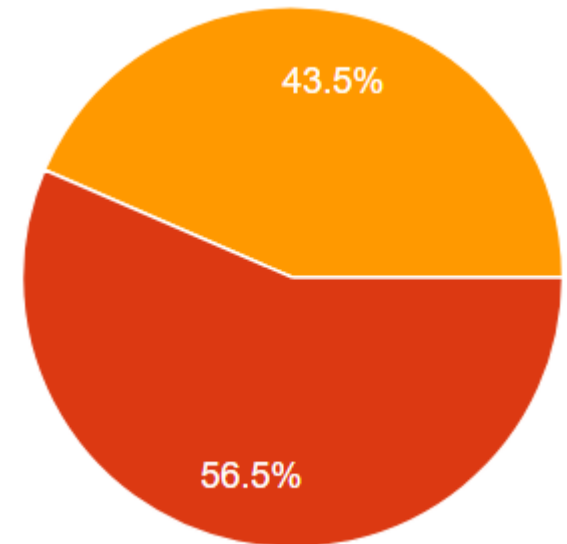
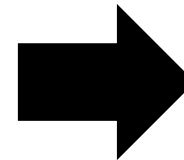
 Helper Selected	 Hinderer Selected	Total
14	2	16

生徒アンケートの結果(参考程度)



特進の平均が88点、進学が平均が90点。
進学より特進の方が優秀だ！

- 全く反論できない
(納得してしまう)
- 本当にそうか？と
疑問を持つことはできる
- 統計的仮説検定の仕組みを
相手に説明し、反駁できる



5段階自己評価

平均(SD)	難易度(1易-5難)	理解度	興味関心
講義	3.3(0.8)	4.1(0.6)	4.3(0.8)
実習	3.3(0.9)	4.1(0.7)	4.4(0.7)
日常におけるデータへの 視点の変化		4.1 (0.9)	



**情報リテラシーは情報をむやみに疑うこと
ではないと実感した。**

統計を学んだ後になるといろいろなことに疑問がでてくる。
例えば私は歌が下手だが、もし小さい頃からピアノを習っていたら歌が上手くなっていたのかなと思って、ピアノを習っていた人とそうでない人でカラオケボックスで国歌を歌って採点結果を比べてみたら違いがあるのではないかと思った。そこで実際に検定して有意になれば、私は自分の子供には小さい頃からピアノを習わせたいと思う。

やっぱり「**統計**」は面白い！



高校生が楽しく学べる
仮説検定と相関・因果

-シミュレーションから始まる“情報科”統計教育-

END