社会科学 (3、5) 2020

高校数学Iから 大学の推測統計学への橋渡し

担当:橘 永久

(Towa Tachibana)

法政経学部、専門:開発経済学

4年前にこの講義を受講した某学部一年生の感想

当時の私のゼミ生が、所属するサークルで聞かされた、とのこと。

何やってんだか、まったく分かんない。



何をやるか(講義内容)

1: 高校数学の記述統計から、大学の(推測)統計学への橋渡し (数学Bの「確率分布と統計的な推測」とほぼ同じ、そしてもう少し高度な内容も)

- 2: 特色として
- a. Association(相関)ではなく、Causality(因果律)の概念・重要性を理解
- b. 哲学的側面の紹介 (オンラインとなったので、 予定よりずっと控えめです)

その他の特色

• 英語参考書を(時々)使用

千葉大学は、

スーパーグローバル(SG)校!

この講義を真面目に受講すると何が分かるか

最大の便益・収穫は

次のような本・雑誌記事の意味が正確に分かり、もしかしたら自分の 分野の学術論文まで読めるようになる!

(というのは過大広告で、正確には「読む時の困難さが、多少減る」)。

まずはコロナ関連から

- ここに画像をうまく貼れませんでした(小さくなりすぎる)。
- 2020年4月27日 毎日新聞夕刊のコピーを今週の資料として掲示すするので、参照してください。

黄色部分に注目

相関か因果か?の議論 (チョコレートの話は、雑誌Newton 2019年4月号 pp.64-65に解説あり)。

中西準子 Nakanishi Junko

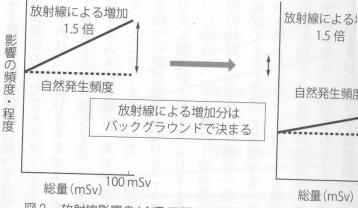
日本評論社

一〇〇ミリ以下推定と科学

3_°

なはずもない。そういう無数の要因を、 のがんの原因がある。それぞれについて、ばくろ量とリスクの関係を疫学で証明することが可能 範囲でのがんのリスクを直接明らかにする力をもたないという合意がある〟と書かれている。 否かは、対象人数、相対リスク、それに交絡因子が関係している。ICRP2007の(A8 長崎の原爆投下の影響ですという形で取り出すことができないということであって、別に、リス 無視できる大きさでもない。いや、 6)にも、√がんリスクの推定に用いる疫学的方法は、およそ一○○ミリシーベルトまでの線量 クがないということではない。疫学研究で、ある原因による影響を定量的に知ることができるか すべき点だ。証明できるぎりぎりの一○○ミリシーベルト被ばくのリスクは、○・五%であり、 と、聞いている方は、リスクはないと言っていると思うのではなかろうか。これは違うし、注意 疫学研究の限界についての言及がなく、「有意な増加は認められない」という表現だけになる ①は、一○○ミリシーベルト以下程度のリスクは、LSS研究の対象人数では、これが広島・ 何らかの推定法を考えて、大まかでもいいからリスクの大きさを推定し、 これは小さいから暫く放置しよう、 これだけなら無視できるかもしれない。しかし、他にも沢山 一つひとつ取りあげて、証明できませんと投げ出してい 一つひとつは小さい





放射線影響の LNT モデルと相対リスクの意義(丹羽太

します。 ますし、

健康な生活を送る努力が大切です。

人生で放射線の影響が現

すでに述べまし

ただ放射

そしてがんが発症

て思わず

印をおされたというものではなく、

肥満を防ぎ禁煙をするとい

った努力でさらに低下

スクは自然に

低下

そうです。

放射線を被ばく

た方々

ます。 では、 を獲得 線がが して、 れるチャ 放射線による単位線量当たりの発がん性突然変異の 放射線より他の要因の寄与の方が大きいことになり また他の要因で四つの突然変異を獲得 がんのリスクは線量の一乗で増加します。 んを増加させる対象は、他の要因で四つの突然変異 した細胞をもつ集団です。 ンスはどんどん少なくなります。 しないほど、その後の そうなると放

相対リス の集団に比例 (一〇〇ミリシ します 1 ので、 IV では一 図2に示したように自然発がん集団を母数として、 五倍) をかけると放射線発がんリスクが出 3 1) 5

の要因で四つの突然変異を獲得

た集団の大きさは

率をかけ算

たものががんの誘発確率となるわけで

した集団に対

射線発がん

鉱山での て禁煙後の相対 通常の電離放射線にかぎらず 仕事をやめた後に、 山やタ ます。たとえばアル リスクにつ 15 コで受け た変異数に応じた乗数で低下 10 年齢の三乗で低下 て計算すると、 ファ線を出すラドンが高い鉱山で働く労働者の肺が その他の要因に だいたい年齢の五乗の逆数で急速に低 します よる発がんでも、 (Tomasek et al., 2008)° するのです 経年的な相対 またタ 1) h ス 0 ク の低下が コに ス クは 0

放射線影響の未然防止のためには、 LNTを考えるしかない

問◆ ŧ, またどの れまでの お話では、 ような場合でも L NTが基本であることは理解できま LNTなのでしょうか? した、 でも 線量が n ほ ど低

丹羽 界的にも まず低線量でLNTが成立するか否かの問題を考えま 0 とも信用があり 以下でのリ ます。 被ばく者の全がんの頻度は スクの増加は統計的に有意ではありません 一〇〇ミリグレ よう。 被ば く者のデ 以上の線量で直線 タ 世



加えて、

機械学習、テキストマイニングなどへの(ほんの) チョットの)導入

AIの基本となっている、機械学習 (machine learning)

Barber, David (2012) Bayesian Reasoning and Machine Learning, Cambridge: Cambridge University Press

前書きでは、

"Machine learning is the study of data-driven methods capable of mimicking, understanding and aiding human and biological information processing tasks."

今週の参考に入れているNatureのマシンラーニング・レビュー論文

• 2ページ目 Box 1が解説しているのは、条件付確率に基づくベイズ の定理

本講義の強調点を、初回に少しだけ説明

因果関係と相関関係の違い



AssociationとCausalityの違い









pixta.jp - 4087489



pixta.jp - 23353882





But (しかし)・・・



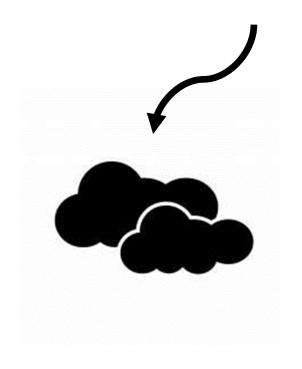


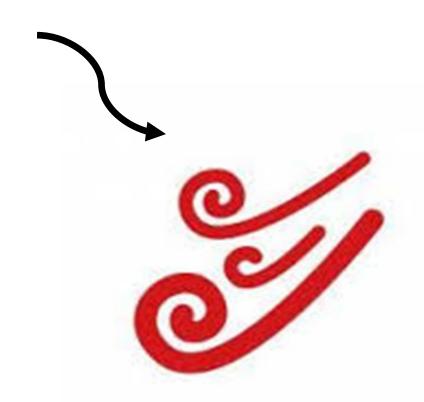




こんな可能性も「ありうる」

北にある冷気(冷たい空気の塊)





いよいよ肝心の

成績評価法

その前に、科目登録をキチンとしてください!!!

成績評価1

 期末試験 80% (今年は、朝に Moodleに問題を出して、当日17時まで に提出してもらう形式を考えています)

2) 宿題一回 20%

試験は、統計・確率の計算問題と統計の思想の理解を問う問題。

成績評価2

参考 (というより老婆心)

統計学は、数学の一分野でもあります。オンライン講義だからと自分に合った教科書も買わずきちんと学ぶこともせずに過ごし、期末試験当日に、直前に買ったばかりの新しい参考書を広げても、大変なことになる可能性が大です。

なんとなく自分の考え(感想)を書けば単位が来る講義では、ありません。



参考: 2019年度の成績

	社会科学5 工学部·看護学部
秀	10 (16.1%)
優	20 (32.3%)
良	19 (30.6%)
可	10 (16.1%)
不可	3 (内、3名未受験) (4.8%)
計	62

普遍科目では、成績分布の目安が大学側から指示されている。

- 1) 秀は%ぐらい、不可は5%以下、等。
- 2) つまりは、相対評価。

実はここにも、統計学の考え方が、いささか安易に適用されている。

真つ当な講義であれば、成績は釣り鐘型(我々が本講義で深く学ぶ正規分布)に従うはずである。

オフィスアワー

•方法を検討中。Moodleのアナウンスメントでお知らせします。

ここから講義本番

• まずは、大学生活の心構えから(できるだけ短くします・・・)

普遍科目の目的(私の理解)

=大学での「学び」への導入

大学での学びの姿勢の最重要項目 3つ

- 1. 「分からない」ということを、あまりに気にしない
- 2. (論理的に)発言する、質問する。
- ある特定の学問(論理)体系にどっぷり浸かる (できれば身につけるように努める)

1. 「分からない」ということをあまり気にしない

大学で学ぶ内容

- 1) 必ずしも決まった答えのあるものばかりではない。
- 2) 聴いただけで分かるような、簡単な内容ばかりでもない

2回、3回と繰り返さないと理解できない内容がザラにある。 1回目では、「何が分からないのか」が分かっただけでも大収穫!

分かり易い講義 の落とし穴

担当教員がとてもとても優秀な場合もあれば、NHKスペシャルを見ていれば済むような表面的な講義の可能性もある・・・



2. 発言する、質問する(メディア講義ということで、質問の受け方を考え中ですが・・・)

大学:

既に分かっていることを学習するのではなく、まだ答えのない問題・新たな問題に取り組む

大卒

ルーティンワークを繰り返すだけではなく、前例のない問題に取り組める人(であってほしい:いわゆる大学=最高学府説)。

論理的・体系的に発言・質問して問題の解決に少しでも貢献する。

Globalization (国際化)

望む・望まないにかかわらず、またコロナで少し後退しても、進行中

世界には

「発言しない人間は死人と同じ」、という文化もある(というより米中を筆頭に、結構多い)。

「世界で働くプロフェショナルが語る」 江川 雅子編 (2014) 東京大学出版会

p.69 上田隆文(元 ILO職員)

『・・・そして発言することです。発言するから存在意義があります。逆に発言しないで会議に行ってメモだけとって帰ってくると「また日本人が情報だけ盗んで帰っている」「あいつら何も貢献しないで奪っていくだけだ」と思われるのが関の山です。』

p.181 樋口泰行(日本マイクロソフト代表執行役社長)

『(日本人が)カーライルの「雄弁は銀、沈黙は金」の後半だけを掲げて、寡黙であることのいいわけをしても、海外では不思議な顔をされるのがオチです。』

で「大学での学び」を終了

• ここからは、統計学の体系にどっぷり浸かります。

・大学での一回目!

・二回目は、工学・看護など皆さんの専門課程で学ぶことになるはずです。

|

本講義の主要な参考書

Agresti, Alan, Christine Franklin and Bernhard Klingenberg (2017), Statistics: The Art and Science of Learning from Data 4th ed., Boston: Pearson Education.

定評のある、高校から大学への橋渡し統計教科書

難点:

値段が高い! 5月5日にアマゾンで見たら、8,360円だった。



(ちなみに)引用文献の示し方

複数の著者がいる場合、First author, Second author, third author,

記法は学問体系ごとに種々あるが、例えば経済学で最も一般的なのは

First Author

Second author 以降

Family name, Given name, Given name Family name, Given name Family name,

Agresti, Alan, Christine Franklin,

図書館では、本講義に関する参考書籍をまとめて開架してくれています。

「授業資料ナビゲーター」

https://alc.chiba-u.jp/pathfinder/index.html

- オンライン講義が正式に決まる前にリストを出したので、いささか哲学寄りですが、参考にしてください。
- なぜか昨年度のナビゲーターもまだ残っており、リンクによってはそちらに飛びます(あまり変わらないけど)
- ・ 千葉大の図書館は、まさに「優れもの」です。自粛期間が終われば、密集しないよう気をつけながら、大いに活用してください。

→【木館ι 坤2陛Λ /117/DVT】

授業資料ナビゲータ(PathFinder)	教養コア:論理コア(論理・哲学・社会)
	2020年度 授業コード:G14D40105
社会科学(3)・(5)	
火4, 木1 橘 永久先生	
キーワード: 推測統計学 因果推論 ——統計学入門:確率と因果推論の哲学	
✓ ←実際に利用した資料にはチェックをしましょう。 ★のついている図書は、授業開始から1年間は本館N棟2階授業資料ナビコーナ	一に配架します。
図書 本には、テーマに関する内容がまとめられています。授業のテーマをつかむには	はまずはここから始めましょう。
□ Statistics: the art and science of learning from data / Alan Ag Klingenberg 4 th ed., Global edition Pearson Education Limited なる → 講義のメイン参考書として使っています。	2018
□ 挫折しない統計学入門:数学苦手意識を克服する/浅野晃著 →講義のサブ参考書です。	★【本館L棟3階大型A 417/STA】 『 オーム社 2017
□ 最速の推計統計:正規分布の徹底攻略/渡邊洋著 コロナ社 → 講義のサブ参考書です。	★【本館L棟3階A 417/ZAS】【電子書籍あり】 2016
	★【本館L棟3階A 417.6/SAI】
参考資料	
□ Rによる統計学独習:データサイエンスの基礎/地道正行著	裳華房 2018 ★【本館L棟3階A 417/RNI】
□ Pythonで学ぶあたらしい統計学の教科書 / 馬場真哉著 翔泳	

IIII

とはいえ、本は

買うもの、使うもの

- 1. 経済学の基本概念の一つ: Opportunity Cost (機会費用)。 スマホで撮ったり、手で写したりしている(馬鹿げた)時間 に、バイトをしたり映画でも観た方がよほど生産的。
- 2. 著作権:他人の知的生産物(本)の半分以上をコピーする のは盗み。
- 3. 後で古本屋に売ろうと綺麗においておいても、売値は買値の10分の一にもならない。徹底的に書き込みして、使い倒した方がお得。

当然、予算によるが、

・できれば入門教科書は3冊あった方が良い。

pp.126 - 127

鎌田 浩毅 (2009)、『一生モノの勉強法』、東洋経済新報社

立花隆(1984)、『「知」のソフトウェア』、講談社。 にも似た記述がありました。

やっと本題:



科学と証拠

統計学とは、科学の問い・仮説に対して、与えられたデータから何が言えるか(=統計的推論(Statistical reasoning))、を導くための手段・方法。

科学の問いの例(p. 28 of Agresti et al. 2017) "Is there bias against women in appointing managers?"

今回の参考資料で言えば、「BCG接種は、コロナウイルス対策として有効か?」

Statistical reasoning (キーワード1):

もう少し柔らかく言い換えると

手持ちのデータ(サンプル sample)から、与えられた問いに対して確かに言えることは何か?

まず、科学とは何か?

・キーワード 2:我々が、通常科学と言っているものを正確に言うと、

Empirical Science (通常の訳語: 経験科学)

狭義の科学・その定義 (理科系)

- 酒井 邦嘉(2006)、『科学者という仕事』、中央公論社, p.19
- •「自然法則の解明」
- ・米沢登美子(2005)、『人物が語る物理入門(上)』、 岩波書店、p. 7 近代科学の骨子
- 1) 対象は森羅万象の自然の営み。
- 2)方法は「実証主義」

最も厳密なFeynman (1963, p.1-1)の定義

Feynman, Richard P., Robert B. Leighton, and Matthew Sands (1963), *Lectures on Physics, Vol.* 1, Reading, Mass: Addison-Wesley

The principle of science, the definition, almost, is the following: The test of all knowledge is experiment. Experiment is the sole judge of scientific ``truth."



Feynman 定義の続き

But what is the source of knowledge? Where do the laws that are to be tested come from? Experiment, itself, helps to produce these laws, in the sense that it gives us hints. But also needed is imagination to create from these hints the great generalizations---to guess at the wonderful, simple, but very strange patterns beneath them all, and then to experiment to check again whether we have made the right guess.



科学の一般的手順

- 1. 論理的な推論、又は(繰り返し)観察に基づいて「仮説」を提示
- 2. 仮説を、実験・繰り返し観察により「実証的に」検証。
- 3. 検証にパスすれば、一応、科学者仲間から命題として受け入れられる。



一応の意味:パラダイムの転換

• 科学史家 トーマス・クーン

The impersonal asocial view of scientific method is inconsistent with history

天動説から地動説へ

Science is not an impersonal and continuous accumulation of the facts drawing ever closer to the truth, but rather, a community and period phenomenon much like art.

50代半ばの担当者は、科学信仰世代

とはいえ、子育てを通じて触れた恐竜研究を通じて、このパラダイム 論を実感することとなった。

・次ページは、筆者が浪人していた時の「科学」誌掲載の、ティラノサウルス像

Newton 2013年 9月号より



1983 年のNewton 掲載のティラノサウルス

過去の Newton でも、当時の考えにしたがって、ティラノサウルスはゴジラのようにえがかれた。イラストは、1983 年10 月号に掲載されたティラノサウルスだ。当時はほとんどの恐竜が、尾をひきずった姿勢でえがかれていた。

それが今や次ページのように・・・

もっと最近なのでは、成獣のティラノサウルスに羽毛を付けているものもある。

土屋健 (2015) 「白亜紀 の生物」





パラダイムの転換

The Structure of Scientific Revolution

Normal science of the period \rightarrow crisis \rightarrow Revolution \rightarrow New Normal Science \rightarrow new crisis



経験科学における仮説の提示

p. 18 of Popper, Karl R. (1959), *The Logic of Scientific Discovery*, New York: Routledge (reprinted version 1992).

『反証可能性(falsifiability)』が必須。

実験で反証する。



Popper p.18 からの引用

'But I shall certainly admit a system as empirical or scientific only if it is capable of being tested by experience. These considerations suggest that not the verifiability but the falsifiability of a system is to be taken as a criterion of demarcation. In other words: I shall not require of a scientific system that it shall be capable of being singled out, once and for all, in a positive sense; but I shall require that its logical form shall be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical scientific system to be refuted by experience.'

仮説の提示法(酒井 p.23)

「お化けは存在しない」

は、科学的な仮説の提示か?

少し先走るが、統計的検定での仮説提示法

帰無仮説 (null hypothesis)

$$\hat{\beta} = 0$$

棄却(reject否定)できて、初めて意味がある。



数学は科学ではない!

• Feynman et al. (1963, p. 3-1)

'Mathematics is not a science from our point of view, in the sense that it is not a natural science the test of validity is not experiment.'

Science Which Is Not a Science...

• Feynman のインタビューをまとめた記事の小見出し。

P. 22 of The Pleasure of Findings Things Out

Science Which Is Not a Science... II

"Social science is an example of a science which is not a science; they don't do things scientifically, they follow the forms—or you gather the data, you do so-and-so and so forth but they don't get any laws, they haven't found out anything."



厳密な実験・厳密に反復可能な観察

経済学などの社会科学 (Social Science) では確かに不可能。

例)教育の経済効果

大学教育がどれほど生産性を高め、生涯所得を向上させるか?

財政逼迫の日本で、重要な問いの一つ(教育に予算を割く必要はあるか?)

どんな実験が必要ですか?



厳密な実験は、実は多くの自然科学でも不可能

• Wilson, Edward O. (1992), The Diversity of Life,(出版社後で・・) p. 170

生物多様性に関する著名な入門書・もはや古典

"Then there might be an answer to the question I am asked most frequently about the diversity of life: if enough species are extinguished, will the ecosystem collapse, and will the extinction of most other species follow soon afterward?

Wilson 続き

The only answer anyone can give is: possibly. By the time, we find out, however, it might be too late. One planet, one experiment."



環境ホルモンのケース(永遠の待ちぼうけ)

Colborn, Theo and Dianne Dumanoski (1997), Our Stolen Future,
 p.196

『奪われし未来』

"For these reasons, those who demand such definitive "proof" before reaching a judgment are certain to be waiting an eternity.

環境ホルモン続き

In the real world, where humans and animals are exposed to contamination by dozens of chemicals that may be working jointly or sometimes in opposition to each other and where timing may be as important as dose, neat cause-and-effect links will remain elusive."

トランプ大統領のような地球温暖化懐疑派

厳密な科学の定義に従っている。

地球において、他の自然要因、火山・太陽活動など全てを一定に保ったままで、人為的温室効果ガスの排出増の効果を計ってみる(=厳密な実験)、などできない。

Feynmanの厳密な「実験」定義に従うと

・厳密には

「教育の経済効果」などの社会科学分野の問いのみならず、

「人為的温室効果ガス排出増による地球温暖化」・「生物多様性の保護」といった地球規模の大問題から、冒頭の「BCGがコロナ対策として有効か」に至るまで

科学的議論はできなくなる。

ガチガチの科学では少し危ないので

The Precautionary Principle

に従うべき、との考え方が一般化。

また最近は、もう少し緩い科学の定義が一般的となっている

・ 手順に従って、追体験(追試)ができるか。

Reproducibility (再現できるか?)

ちなみに私(橘)は、5月3日8時に神 の啓示を受けた!

こうした科学の定義に従い、反証可能な仮説が提示されたとする

問いの例(p. 29 of Agresti et al. 2017)

"Does a low-carbohydrate diet result in weight loss?"

効果は無い、という仮説を検証すればよい(我らが 最終目標:帰無仮説の検定を想起しよう)。 先ほど、統計学とはいかなる学問かを描写

統計学とは、科学の問い・仮説に対 して、与えられたデータから何が言 えるか(一統計的推論(Statistical reasoning))、を導くための手段・方 法。

統計学 (Statistics)の定義は?

参考書 Agresti et al. (2017) P.30

"Statistics is the art and science of designing studies and analyzing the data that those studies produce. Its ultimate goal is translating data into knowledge and understanding of the world around us. In short ..."

他の教科書での統計学の定義 1

田中勝人 (1998)『統計学』 新世社 p.i はしがき

「統計学は、端的にいえば、統計データを分析するための方法論的科学である。ここで、統計データとは統計分析の対象となるデータを意味する。・・・統計学はデータなしでは成立しない学問である。

ところで方法論としての統計学は、データの収集から始まって統計的結論を得るまでのプロセスを段階的に分類することにより、次の4つの方法にまとめることができる。

- 1) 標本調査、実験計画
- 2) 記述統計
- 3) 推定、検定
- 4) 決定、予測

英文参考書p.31で、対応する部分は

Design

Description

• Inference



他の教科書の統計学定義

和達三樹・十河清(1993)『キーポイント 確率・統計』 岩波書店 p.2 「…統計学では、多くの事象を含む集団からいくつかを標本として取り出し、その標本をもとにして集団の性質を推測する。」

科学雑誌 Newton 2013年12月号 特集「統計の威力」 p.48 「統計学とは、膨大な母集団すべてを調べるかわりに、一部のデータ(標本)をとり、そこから母集団の特徴を効率よく推定する技術である。」

Bulmer (1967)、縄田和満 (2013) 『東京大学工学教程 確率・統計」 明確な定義なし! (そんなこと分り切っている、又は使う時にはどうでも良い から、サッサと本題に行こう!という感じか)

高校教科書:数研出版 新編 数学 データの分析:こちらも定義なし

お薦めの簡単な定義

参考書 p.30 の In short 以降、 Statistics is the art and science of learning from data (~sample (標本)).

気を付ける言葉

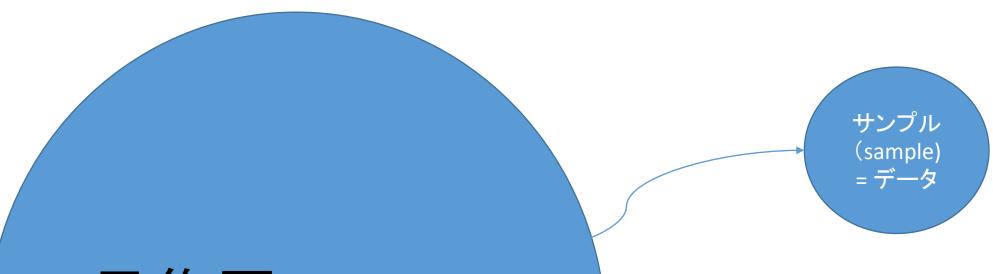
Art: 芸術というより、技能という訳の方が良い。

統計学のもう一つの重要な定義

Quantifying uncertainty: よって確率(probability)が重要な役割を果たす

社会科学5履修生に、少なくとも10年は覚えておいて欲しい概念が次の図





母集団 (population)

統計学の定義の裏にあるもの:

1. 世界観の変遷が背景にある

デイヴィッド・サルツブルグ (2001, 翻訳2006)『統計学を拓いた異才たち』 日本経済新聞社、前書きと第2章

統計学が科学の世界で市民権を得る

=ニュートン力学に代表される決定論世界 観(時計仕掛けの世界:初期条件と正確な 公式があれば、全宇宙の運動を予測でき る)の崩壊

ピエール・シモン・ラプラス (1749-1827)

計測の揺らぎは測定の誤差であり、計測が精緻化すれば消え去る(ラプラスの魔)、と考えた(=科学が不完全だから確率や統計学が必要、という立場と言える)。

誤差は、誤差関数としてまとめてしまう。すると誤差は、特定の分布で表現できた(=正規分布)。

因みに杉田玄白の『蘭学事始』が1815年

Karl Pearson (1857-1936)による革命

・実験の観測値は、偶然の産物に過ぎない。(どこぞで蝶が羽ばたいたことによる空気の僅かな揺れ・温度の変化・磁気の微妙な変化で正確な繰り返しはできない)。

サルツブルグ p.19

「実験結果から得られる値は散らばっており、その中の一つが正しいのではなく、正しい値にできるだけ近い推定値を得るためにすべてのデータを用いることができるのである。」

観測値は確率分布に従っている

Karl Pearson以前は

科学の対象は、実在し知覚できるものと考えられていた惑星の運行、特定の動物の血流等。

ピアソン: 観測値は偶然に左右されており、実在するものは惑星の運行・動物の血流など科学対象の確率分布、と見方を変えた。

"The `laws of Nature' are only constructs of our minds; none of them can be asserted to be true or false, they are good in so far as they give good fits to our observations of Nature, and are liable to be replaced by a better 'fit'..."

ちょっと脇道の情報を足しておくと

カール・ピアソンの統計学の考え方を、「記述統計学」といい、現在主流の「推測統計学」とは異なる。

現在の主流は、

自然(究極のLaw of Nature)そのものもランダム(なのかもしれない)

- ・ 先ほどの縄田の教科書 p.1 「はじめに」の一部
- 「・・・量子力学などでは、電子などの挙動はすべて不確実性を伴っている。このような不確定性(注: 不確実ではない) やばらつきに対応するためには、数学的な道具として、どうしても確率や確率分布・確率変数の基礎的な知識が必要である。」

ピアソンの見方は、現在の'パラダイム'

サルツブルグ p.30

「...(多くの人々は)自然は基本的にランダムであり、ただ一つの真理は分布関数のなかにあると信じている。」

クーンの科学史分析(The Structure of Scientific Revolution)が示したように、こうした世界観(パラダイム)は、将来ひっくり返る可能性はある。

「神はサイコロを振り賜わず」

が主流に返り咲く可能性もある。

我々は、現在のパラダイムのなかで、一般 的な統計学を学んでいく。

ちなみに、「一般的な」も変わり得る

- 1) 私が大学生だった35年ほど前は、「ベイズ統計学」は異端だった。人工知能ブームもあり、今や生協書籍部の統計学の棚には、ベイズと付いた本がいっぱいある。
- 2) そもそも私が大学生だったころに、生協書籍部に、統計学で棚がいくつ かある、などということは無かった(統計学の人気化)。
- 3) 本講義が最終到達点に掲げる有意性に基づく統計的仮説検定だが、 近年、非常に評判が悪くなっている(Agresti et al. 2018, Ch. 9 Section 5)。
- 4) それでも、まずは現在のパラダイムのもとで、「一般的な」統計学を学ぶ 意義は大きい、と私は信じている。

現在一般的な統計学の定義の裏にあるもの:

- 2. 知りたいもの(母集団)について知るのは大変すぎる
- 先ほどの縄田の教科書 p.1 はじめに
- 「・・・データ分析において対象とする集団全体を母集団とよぶが、母集団全体について知ることは、ほとんどの場合困難であり、多くの場合、母集団からその一部を選び出し(注:標本)、選び出されれた集団について調査を行い、母集団について推定するということが行われる。」

人工知能の発展で、膨大なデータの蓄積が可能に=ビッグデータ 母集団そのものを分析できる時代がすぐそこかも?

統計的手法の手順

p.8(ページ下の数字)統計学の定義のすぐ下

1) Formulate a statistical question

2) Collect data

3) Analyze data

4) Interpret results

p. 10 のDesign (設計)にあたる。

参考書 p.34 Population と Sample

Population (母集団): Formulate a statistical question

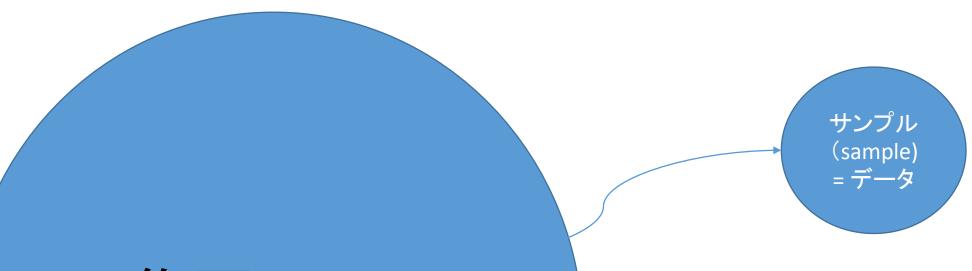
"...the total set of subjects in which we are interested."

Sample (標本): Collect data

"...the subset of the population for whom we have data."

多くの場合、我々はSampleしか持っていない(母集団は、'いささか' 漠然としている)。しかし、興味は母集団にある。





母集団 (population)

Inference (推定)

p.34 Inference (推定) in Statistics (統計学): Analyze data

標本 (sample) から情報を効率的に搾り取って(exploit the data)、母集団の特性(パラメーター: parameters)を科学的に予測する。

母集団は、'いささか' 漠然としていることもある

Agresti et al. (2017), p.34

Sometimes the population is well-defined, but sometimes it is a hypothetical set of people that is not enumerated.

母集団とサンプルの例 (Causalityも絡めて)

私の専門分野:経済学からの例。以下は、なんとなく分かってもらえれば十分。

Carrell, Scott E. and James E. West (2010), "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy*, 118(3), 409—432.

THE statistical question in this paper

Major Question:

大学教員の質(学位、教育経験、学生からの評判?)が、大学生の能力向上にどれほどの影響を与えるか?

Minor Question: そもそも大学教員の教育の質を、いかにして測るか?



Carrell and West (2010):続き

Population(母集団):

All the lecturers (professors, associate professors, assistant professors) and undergrads in the U.S. colleges and universities (time not specified).

アメリカの大学生と教員"全般"

(アメリカの大学教育全般に関して、なにか言いたい、でもいつの時代の?)

Sample(標本):

Lecturers and undergrads of the U.S. Air Force Academy (空軍士官学校)

91 lecturers, 10,534 students from the fall semester of 2000 to through the spring semester of 2007.

このstatistical questionの何が難しいか?

ある年に限れば、全米の大学生と大学教員を全員、即ち母集団を確定することは理論上可能。

しかし統計学の標準手法に従い、母集団のサイズに応じたサンプルサイズを確保しても、もっと極端なことを言えば、母集団全体に対する全数調査(センサス)を実施しても、このstatistical questionには答えられない。



米国: Measures (指標) of teacher quality

• Contemporaneous student performance on standardized achievement tests (主に、primary and secondary schools)

• Scores on student evaluations of professors (主に大学)

• Relative student performance in follow-on classes (この論文のselling point)

このstatistical questionの何が難しいか? 2

甘い教員や期末試験や資格試験対策を集中的に教えてくれる教員が「分かりやすい」・「丁寧」・「実践的」として学生が集中し、さらに学生評価で高い評価を得がち

Sample self-selection biasという

試験対策が教育か???学生の直近の成績は急上昇するかもしれないが、本質を理解できずに将来伸びない可能性・・・。能力向上をどこで測るか、という問題。

逆に、

出来る学生ややる気のある学生のみが、「厳しい」という評判の教員 に集中するかも:

厳しいと評判の教員の講義履修者が出来る学生になるのは、もともと履修者の出来がよかったからだけかも・・・。

これも Sample self-selection bias

Carrell and West (2010)は特殊なデータ(サンプル)をフルに活用して、この2つの難題を回避した。

Collect data

自ら実験や調査をしてsampleを作る場合もある(標本調査、実験計画)。

Carrell and West (2010)は、既存のデータの中で、自らのresearch question に挑戦するのに適切なデータ(サンプル)を探し出した。

The United States Air Force Academyの学生成績と、教員の評価記録。



なぜこのサンプルが適切か

1年次の必修科目が標準化されており(教える内容、試験問題が同じ)、かつ

Students are randomly assigned to professors over a wide-variety of standardized core course.

しかも、2年生以降の必修科目の成績が分かる!

Sample self-selectionが無いので、causality(異なる教員 次年度以降の学生の成績の伸び)が分かる!

Carrell and West (2010)の結果だけお伝えしておくと

・学生アンケートで高い評価を得た教員の講義をとった学生は、1年目の成績は良かったものの、2年次以降の共通科目で良い成績が取れなかった。

- ・教育歴は長いが学生アンケートの評価は低い教員の講義に割り当てられた学生は、1年目の成績は悪いものの、2年次以降の共通科目では顕著に(統計的に有意に)良い成績を修めていた。
- •解釈1 deep learning (基本や勉強法を、教えている・学んでいる)
- 解釈2 ????

TABLE 5
PROFESSOR OBSERVABLE CHARACTERISTICS AND STUDENT EVALUATIONS OF PROFESSORS

	λ_{j1}^1 (1)	λ_{j1}^2 (2)
	A. Professor Observable Attributes	
Associate/full professor	69*	.70*
	(.41)	(.40)
Terminal degree	28	.38
	(.27)	(.27)
Greater than 3 years' teaching experience	79***	.66**
	(.29)	(.29)
	B. Student Evaluation Scores	
Instructor's ability to provide clear, well-organized in-	.51***	46**
struction was	(.19)	(.20)
Value of questions and problems raised by instructor was	.70***	59**
	(.24)	(.25)
Instructor's knowledge of course material was	.56**	44*
	(.24)	(.24)
The course as a whole was	.49**	39*
	(.23)	(.23)
Amount you learned in the course was	.59**	47*
	(.23)	(.24)
The instructor's effectiveness in facilitating my learning	.54***	45**
in the course was	(.20)	(.20)

Note.—Each row by column represents a separate regression in which the dependent variable is the Bayesian shrinkage estimates of the corresponding professor random effects estimated in eq. (1). In all specifications the Bayesian shrinkage estimates were scaled to have a mean of zero and a variance of one. Panel A shows results for modal rank and mean years of teaching experience. Panel B shows results for sample career averages on student evaluations.

^{*} Significant at the .10 level.

^{**} Significant at the .05 level.

^{***} Significant at the .01 level.