data-ppf.github.io Feb 25 2020

lecture 6 of 14: data gets real: mathematical baptism

chris wiggins + matt jones, Columbia

logistics/housecleaning

logistics/housecleaning

- see FAQ for new schedule for HW, etc

logistics/housecleaning

- see FAQ for new schedule for HW, etc
- Op-Ed assigning early via `#general` in Slack

student reactions

student reactions: some data on people

```
175 fisher/fisher's
 72 neyman
 46 pearson
 42 guinness
 39 gosset
 22 empire
 15 joan
  9 gigerenzer
  3 quetelet
```

student reactions: statisticians are people too

petty egoism and personal politics

Everyone likes charming turn of the century English scientists... bicycle rides [and] Irish barley farmers. [But] what was happening was a corporations attempt to standardize its product with the best methods modern science had available.

[the fact that] Pearson would not publish Fisher's work because Fisher pointed out errors in Pearson's work seems counter-intuitive to the entire enterprise of scientific discovery

student reactions: hard/more stats

I found these readings somewhat more difficult to access than those in previous weeks. I have yet to take a statistics class, and thus had some trouble grasping the key mathematical terminology throughout the texts.

▶ today is the mathiest lecture

From my classes in mathematics and statistics, I always thought of that problems and methods to achieve a solution were binary - either right or wrong.

student reactions: art+science

"The choice of the test statistic, and of null hypothesis worth testing, remains, for Fisher, an art"

I also enjoyed reading the Guinness story, and could not overlook the influence that capitalism and corporate interest had in advancing statistical development. It is clear that such influence has not wavered in the modern "big data" and statistics arena.

I was starting to worry that eugenics had driven all of the innovations in this field, so I am glad to hear that a nobler cause (in my biased opinion) was at play as well: beer.

themes for today

- early trading zone: industry, agriculture, and mathematics (itself localized to academia)

themes for today

- early trading zone: industry, agriculture, and mathematics (itself localized to academia)
- mathematical-objectivity as truthmaking culminates in algorithmic-objectivity as truth-making

themes for today

- early trading zone: industry, agriculture, and mathematics (itself localized to academia)
- mathematical-objectivity as truthmaking culminates in algorithmic-objectivity as truth-making
- art vs scale (NB: art requires artists, scale requires algorithms)

guiding questions every week:

- Scientific and mathematical development

guiding questions every week:

- ▶ Scientific and mathematical development
- ▶ Technologies and engineering (not so much in "part 1", more in parts 2/3 of class)

guiding questions every week:

- ▶ Scientific and mathematical development
- ▶ Technologies and engineering (not so much in "part 1", more in parts 2/3 of class)
- ▶ Driving forces: money, prestige, resources, Imperial competition... and "truth!"

guiding questions every week:

- ▶ Scientific and mathematical development
- ▶ Technologies and engineering (not so much in "part 1", more in parts 2/3 of class)
- ▶ Driving forces: money, prestige, resources, Imperial competition... and "truth!"
- ▶ how did new capabilities rearrange power? (who can now do what, from what, to whom?) (rights, harms, justice)

context: people, ideas, things

people

(see also timeline)

1. 1701-1761 Thomas Bayes

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson
5. 1876-1937 'Student,' aka William Sealy Gosset

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson
5. 1876-1937 'Student,' aka William Sealy Gosset
6. 1890-1962 Ronald Fisher

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson
5. 1876-1937 'Student,' aka William Sealy Gosset
6. 1890-1962 Ronald Fisher
7. 1894-1981 Jerzy Neyman

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson
5. 1876-1937 'Student,' aka William Sealy Gosset
6. 1890-1962 Ronald Fisher
7. 1894-1981 Jerzy Neyman
8. 1895-1980 Egon Pearson

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson
5. 1876-1937 'Student,' aka William Sealy Gosset
6. 1890-1962 Ronald Fisher
7. 1894-1981 Jerzy Neyman
8. 1895-1980 Egon Pearson
9. 1902-1950 Abraham Wald

people

(see also timeline)

1. 1701-1761 Thomas Bayes
2. 1777-1855 Gauss
3. 1822-1911 Francis Galton
4. 1857-1936 Karl Pearson
5. 1876-1937 'Student,' aka William Sealy Gosset
6. 1890-1962 Ronald Fisher
7. 1894-1981 Jerzy Neyman
8. 1895-1980 Egon Pearson
9. 1902-1950 Abraham Wald
10. 1903-1987 Andrey Kolmogorov

Bayes

Figure 1: Reverend Bayes. We will need him when we go to war (next week)

Bayes and God

- in his essay: 0

The True Title of Bayes's Essay

Stephen M. Stigler

(Submitted on 1 Oct 2013)

New evidence is presented that Richard Price gave Thomas Bayes's famous essay a very different title from the commonly reported one. It is argued that this implies Price almost surely and Bayes not improbably embarked upon this work seeking a defensive tool to combat David Hume on an issue in theology.

Figure 2: Stigler on Bayes on God

Bayes and God

- in his essay: 0
- context:

The True Title of Bayes's Essay

Stephen M. Stigler

(Submitted on 1 Oct 2013)

New evidence is presented that Richard Price gave Thomas
Bayes's famous essay a very different title from the commonly
reported one. It is argued that this implies Price almost surely
and Bayes not improbably embarked upon this work seeking a
defensive tool to combat David Hume on an issue in theology.

Figure 2: Stigler on Bayes on God

ideas, philosophical

- likelihood

ideas, philosophical

- likelihood
- Bayesian probability

ideas, philosophical

- likelihood
- Bayesian probability
- prior

ideas, philosophical

- likelihood
- Bayesian probability
- prior
  - Fisher willing as long as prior itself is known via sampling

ideas, philosophical

- likelihood
- Bayesian probability
- prior
  - Fisher willing as long as prior itself is known via sampling
  - else: "subjective. . . taste"

ideas, philosophical

- likelihood
- Bayesian probability
- prior
  - Fisher willing as long as prior itself is known via sampling
  - else: "subjective. . . taste"
- significance

ideas, philosophical

- likelihood
- Bayesian probability
- prior
  - Fisher willing as long as prior itself is known via sampling
  - else: "subjective. . . taste"
- significance
- confidence intervals

ideas, philosophical

- likelihood
- Bayesian probability
- prior
  - Fisher willing as long as prior itself is known via sampling
  - else: "subjective...taste"
- significance
- confidence intervals
- statistical null vs. substantive null

likelihood, prior, and other probabilities: example 1/2

- consider island w/10,000 people

likelihood, prior, and other probabilities: example 1/2

- consider island w/10,000 people
- "perfect" test assesses 1% have disease

likelihood, prior, and other probabilities: example 1/2

- consider island w/10,000 people
- "perfect" test assesses 1% have disease
- records lost!

likelihood, prior, and other probabilities: example 1/2

- consider island w/10,000 people
- "perfect" test assesses 1% have disease
- records lost!
- new, 99% accurate test for disease records all people

likelihood, prior, and other probabilities: example 1/2

- ▶ consider island w/10,000 people
- ▶ "perfect" test assesses 1% have disease
- ▶ records lost!
- ▶ new, 99% accurate test for disease records all people
- ▶ if you test "positive" in test, probability of being sick?

	s	h	
y			198
n			9,802
—	—	—	—
	100	9,900	10,000

	s	h	
y	99	99	198
n			9,802
—	—	—	—
	100	9,900	10,000

likelihood, prior, and other probabilities: example 2/2

	s	h	
y	99	99	198
n	1		9,802
—	—	—	—
	100	9,900	10,000

	s	h	
y	99	99	198
n	1	9,801	9,802
—	—	—	—
	100	9,900	10,000

	s	h	
y	TP	FP	198
n	FN	TN	9,900
—	—	—	—
	100	9,900	10,000

likelihood, prior, and other probabilities: example 2/2

	s	h	
y	TPR=99/100	FPR=99/9900	
n	FNR=1/100	TNR=9801/9900	
—	—	—	—
	B=100/10,000	1-B=9,900/10,000	

- note "type 1" and "type 2" errors: FNR=$p(n|s)$ and FPR=$p(y|h)$

likelihood, prior, and other probabilities: example 2/2

	s	h	
y	TPR=99/100	FPR=99/9900	
n	FNR=1/100	TNR=9801/9900	
—	—	—	—
	B=100/10,000	1-B=9,900/10,000	

- note "type 1" and "type 2" errors: FNR=$p(n|s)$ and FPR=$p(y|h)$
- none of which is what you want: $p(s|y)$

likelihood, prior, and other probabilities: example 2/2

	s	h	
y	TPR=99/100	FPR=99/9900	
n	FNR=1/100	TNR=9801/9900	
—	—	—	—
	B=100/10,000	1-B=9,900/10,000	

- note "type 1" and "type 2" errors: $FNR=p(n|s)$ and $FPR=p(y|h)$
- none of which is what you want: $p(s|y)$
- $p(s|y) = 99/(198) = 99/(99 + 99) = 1/2$

likelihood, prior, and other probabilities: example 2/2

	s	h	
y	TPR=99/100	FPR=99/9900	
n	FNR=1/100	TNR=9801/9900	
—	—	—	—
	B=100/10,000	1-B=9,900/10,000	

- note "type 1" and "type 2" errors: $FNR=p(n|s)$ and $FPR=p(y|h)$
- none of which is what you want: $p(s|y)$
- $p(s|y) = 99/(198) = 99/(99 + 99) = 1/2$
- $p(h|y) = \frac{p(y|h)p(h)}{p(y|h)p(h)+p(y|s)p(s)} = \frac{FPR \cdot (1-B)}{FPR \cdot (1-B)+(1-FNR) \cdot B}$

likelihood, prior, and other probabilities: example 2/2

	s	h
y	TPR=99/100	FPR=99/9900
n	FNR=1/100	TNR=9801/9900
	B=100/10,000	1-B=9,900/10,000

- note "type 1" and "type 2" errors: $FNR=p(n|s)$ and $FPR=p(y|h)$
- none of which is what you want: $p(s|y)$
- $p(s|y) = 99/(198) = 99/(99 + 99) = 1/2$
- $p(h|y) = \frac{p(y|h)p(h)}{p(y|h)p(h)+p(y|s)p(s)} = \frac{FPR \cdot (1-B)}{FPR \cdot (1-B)+(1-FNR) \cdot B}$
  - (simplifies to 1/2 if $FPR = FNR = B$)

likelihood, prior, and other probabilities: example 2/2

	s	h
y	TPR=99/100	FPR=99/9900
n	FNR=1/100	TNR=9801/9900
—	—	—
	B=100/10,000	1-B=9,900/10,000

- note "type 1" and "type 2" errors: $FNR=p(n|s)$ and $FPR=p(y|h)$
- none of which is what you want: $p(s|y)$
- $p(s|y) = 99/(198) = 99/(99 + 99) = 1/2$
- $p(h|y) = \frac{p(y|h)p(h)}{p(y|h)p(h)+p(y|s)p(s)} = \frac{FPR\cdot(1-B)}{FPR\cdot(1-B)+(1-FNR)\cdot B}$
  - (simplifies to 1/2 if $FPR = FNR = B$)
    - (e.g., $FPR = FNR = B = .01$)

ideas, mathematical

- null hypothesis

ideas, mathematical

- null hypothesis
- $p(d|h_0)$

ideas, mathematical

- null hypothesis
- $p(d|h_0)$
- $p(d|h_1)$

ideas, mathematical

- null hypothesis
- $p(d|h_0)$
- $p(d|h_1)$
- $p(h_0)$

ideas, mathematical

- null hypothesis
- $p(d|h_0)$
- $p(d|h_1)$
- $p(h_0)$
- $p(h_0|d)$ is unspeakable (in both frameworks)

places

- interwar Poland

places

- interwar Poland
- US

places

- interwar Poland
- US
- Berkeley statistical lab

places

- interwar Poland
- US
- Berkeley statistical lab
- Berkeley statistical department

places

- interwar Poland
- US
- Berkeley statistical lab
- Berkeley statistical department
- RAF: Rothamsted->UCL (eugenics)->Cambridge->Australia

places

- interwar Poland
- US
- Berkeley statistical lab
- Berkeley statistical department
- RAF: Rothamsted->UCL (eugenics)->Cambridge->Australia
- Russia

Readings: (Joan) Fisher, EoC, (R.A.) Fisher, Neyman, Phillips

Context on Guinness in 1900 or so

- 1886 IPO

  In 1886, Arthur Guinness' grandson Edward sold his 65% stake in the company in an initial public offering for 6 million pounds, valuing the company at about 9.23 million pounds. In today's dollars, that valuation represents about $321 billion. Guinness stock soared 60% on its first day of trading.

J. Fisher Box

Origin of name "Student"

- He spent '06-'07 w/ Pearson and

J. Fisher Box

Origin of name "Student"

- He spent '06-'07 w/ Pearson and
- Managing director decided "that such publication might be made without the brewers' names appearing. They would be merely designated"Pupil" or "Student."

J. Fisher Box

agriculture, industry, and eugenics

- ▶ The brewers "had become keen geneticists and breeders since the rediscovery in 1900 of Mendel's paper on hybridization"

J. Fisher Box

agriculture, industry, and eugenics

- ▶ The brewers "had become keen geneticists and breeders since the rediscovery in 1900 of Mendel's paper on hybridization"

- ▶ Fisher formed Cambridge Eugenics Society as sophomore: *"The widespread, fruitful, and successful races of the future belong to the dominant nations of to-day"*

Figure 3: Gosset

From astronomical error to beer! Gosset's internal report for Guinness in 1904

EoC

Eoc's taxonomy:

"how to draw conclusions from statistical data":

1. Fisher

Eoc's taxonomy:

"how to draw conclusions from statistical data":

1. Fisher
2. Neyman + E. Pearson

Eoc's taxonomy:

"how to draw conclusions from statistical data":

1. Fisher
2. Neyman + E. Pearson
3. Bayesian

Eoc's taxonomy:

"how to draw conclusions from statistical data":

1. Fisher
2. Neyman + E. Pearson
3. Bayesian

- ▶ "systematic"

Eoc's taxonomy:

"how to draw conclusions from statistical data":

1. Fisher
2. Neyman + E. Pearson
3. Bayesian

- ▶ "systematic"
- ▶ general

Eoc's taxonomy:

"how to draw conclusions from statistical data":

1. Fisher
2. Neyman + E. Pearson
3. Bayesian

- ▶ "systematic"
- ▶ general
- ▶ [different approaches for different uses? – cw]

getting Bayesian, decision theory, and homo economicus

- ▶ see EoC example:

  Imagine a manufacturer... diameter is 8 millimeters (H_1) or 10 millimeters (H_2) she knows standard deviation ... sample statistic [is] mean diameter for each of the two hypotheses. accept H_1, and reject H_2 [or] stop the production and look for the cause of the apparent malfunctioning.

getting Bayesian, decision theory, and homo economicus

- see EoC example:

  Imagine a manufacturer... diameter is 8 millimeters (H_1) or 10 millimeters (H_2) she knows standard deviation ... sample statistic [is] mean diameter for each of the two hypotheses. accept H_1, and reject H_2 [or] stop the production and look for the cause of the apparent malfunctioning.

- $\mu(a|d) = \sum_h R(a,h)p(h|d) = \frac{1}{p(d)} \sum_h R(a,h)p(d|h)p(h)$
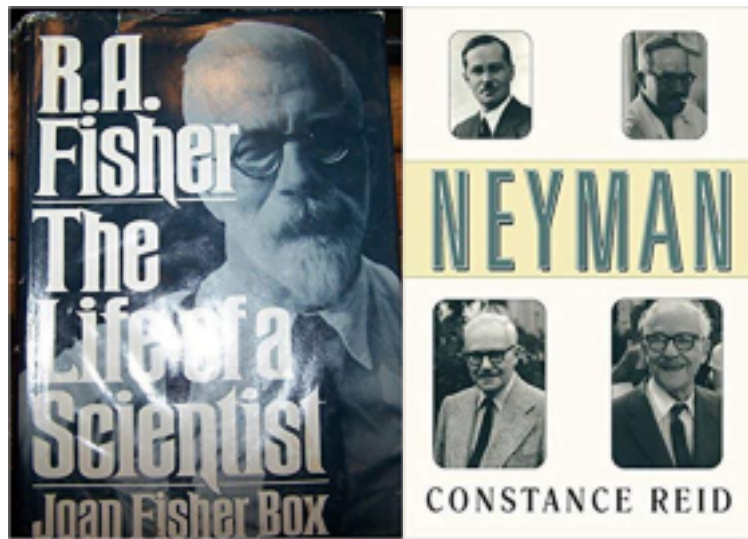
mathbattle: Neyman on 50 years of Fisher vs. Neyman

"inductive reasoning" and "inductive behavior". After a conscientious effort to find the exact meaning of [inductive reasoning], I [conclude] the term is empty

"inductive behavior" means adjust[ing] actions to noticed frequencies of events... mathematical statistics is ...rules of behavior based on observable random variables (for example: rejection or acceptance of hypotheses, methods of estimation, etc.) and, in ...the search for rules (or "strategies") that satisfy some pre-assigned conditions (most powerful tests, unbiased minimum variance estimators, maximin ordering policies, etc.). [Fisher views it as] "geared to technological performance", such as Russian and American [but] inconsistent with the process of drawing correct conclusions in Natural Science.

Figure 4: Fisher-Mackenzie crop experimental design

A plan for a controlled agricultural experiment, from: Fisher and MacKenzie, "Studies in Crop Variation. II. The Manural Response of Different Potato Varieties," *J Ag Sci* (1923)

Fisher: big ag @ Rothamsted Experimental Station 2/2

- large stockpile of older experimental data

Figure 5: "millionaire calculator"

Millionaire calculator purchased for Fisher £££££$$$$ tech of the day!

Fisher: big ag @ Rothamsted Experimental Station 2/2

- ► large stockpile of older experimental data
- ► large operation for the transformation of data

Figure 5: "millionaire calculator"

Millionaire calculator purchased for Fisher £££££$$$ tech of the day!

Fisher: big ag @ Rothamsted Experimental Station 2/2

- ▶ large stockpile of older experimental data
- ▶ large operation for the transformation of data
- ▶ pre-WWII "compute"

Figure 5: "millionaire calculator"

Millionaire calculator purchased for Fisher £££€$$$$ tech of the day!

Fisher: experiment advanced maths

"Daily contact with the statistical problems which present themselves to the laboratory worker has stimulated purely mathematical researches upon which are based the methods here presented."

Fisher: yet stats inadequate for experiment

"the traditional machinery of statistical processes is wholly unsuited to the needs of practical research. Not only does it take a cannon to shoot a sparrow, but it misses the sparrow!"

"The elaborate mechanism built on the theory of infinitely large samples is not accurate enough for simple laboratory data." (1925)

Fisher: tests of significance

probability that null hypothesis would have produced experimental data

No universal threshold of significance!

> *"it is usual and convenient for experimenters to take 5 per cent. as a standard level of significance, in the sense that they are prepared to ignore ail results which fail to reach this standard."*

Fisher: insistence on randomization

Randomization necessary so that "the test of significance may be guaranteed against corruption by the causes of disturbance which have not been eliminated."

Randomization "relieves the experimenter from the anxiety of considering and estimating the magnitude of the innumerable causes by which his data may be disturbed."

Fisher: experiment as conduit to liberty

Science –> truth –> liberty

> *"The liberation of the human intellect must, however, remain incomplete so long as it is free only to work out the consequences of a prescribed body of dogmatic data, and is denied the access to unsuspected truths, which only direct observation can give."*

Neyman (b. 1894): challenger to Fisher (b. 1890)

Neyman: Polish context

- Probability applied to the experimental agricultural work in newly–and temporarily–independent Poland in the 1920s.

Neyman: Polish context

- Probability applied to the experimental agricultural work in newly–and temporarily–independent Poland in the 1920s.

- Part of what historian Theodora Dryer calls "a dynamic movement to imagine sovereign Poland as a modern and prosperous agrarian nation state" through experimentation and most sophisticated theoretical mathematics of the day.

Neyman: science ≠ truth

The problem with most hypothesis testing: most people thought it was about truth. It wasn't.

> *Without hoping to know whether each separate hypothesis is true or false, we may search for rules to govern our behaviour with regard to them, in following which we insure that, in the long run of experience, we shall not be too often wrong.*

- ▶ Neyman and Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," 291

Neyman: science $\neq$ truth

> "Deciding to affirm doesn't mean knowing or even
> believing." Rather, "it's an act of will preceded by some
> experience and deductive reasoning, just as one takes out
> life insurance, which we do even if expect to live for a long
> time."

▶ Neyman, "L'Estimation statistique" (1938)

Fisher's response: unhappiness with abandonment of truth

> *To one brought up in the free intellectual atmosphere of an earlier time there is something rather horrifying in the ideological movement represented by the doctrine that reasoning, properly speaking, cannot be applied to empirical data to lead to inferences valid in the real world.*

► Fisher, Statistical Methods and Scientific Inference, 7.

Fisher's response: Cost $\neq$ science

> *in inductive inference we introduce no cost functions for faulty judgments, for it is recognized in scientific research that the attainment of, or failure to attain to, a particular scientific advance this year rather than later, has consequences, both to the research programme, and to advantageous applications of scientific knowledge, which cannot be foreseen . . . . We make no attempt to evaluate these consequences, and do not assume that they are capable of evaluation in any currency.*

▶ Fisher, "Statistical Methods and Scientific Induction," 1955

Phillips on synthesis/modern impact, e.g., "algos" and FDA / drug approval

FDA/drug approval: a sidebar 2/3

> *In 1960, the secretary of the American Pharmaceutical Association painted a dystopian hellscape that automation and government regulation would soon bring about: "The electronic brains, which have been government programmed, will prescribe the 'correct' drug." The secretary of the pharmaceutical association argued that competition among drug makers would be no more: "There is no problem of competition or concentration since each pharmaceutical manufacturer is assigned the products it can produce, the quality specifications it must not exceed, the price it must charge." . . . Algorithms will subordinate the clinical expertise of trustworthy physicians and pharmacists.*

- ▶ M Jones, unpublished, citations from Tobbell, Pills, Power, and Policy ch. 4.

FDA/drug approval: a sidebar 2/3

> *With these dramatic transformations of the [1962 Kefauver-Harris amendment in response to Thalidomide scandal], the randomized controlled trial became the benchmark for gauging the efficacy of medications, to become the gold standard for authorization of drugs and the documentation of their side effects.*

- ▶ M Jones, unpublished, see also act, scandal

FDA/drug approval: a sidebar 3/3

Figure 1. Funding Sources for Randomized, Controlled Trials (RCTs).
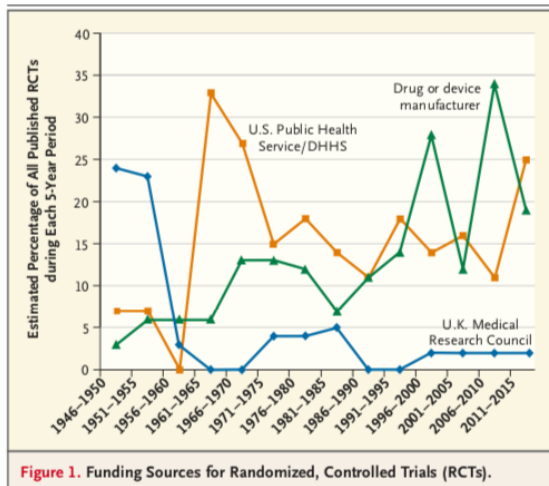
Figure 6: percent of RCT from US gov, UK gov, industry

modern impact 1/2

- ▶ FDA / drug approval

modern impact 1/2

- ▶ FDA / drug approval
- ▶ ethics of the above, power of statistics v. expertise

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo
  - ethics of harmful drugs

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo
  - ethics of harmful drugs
  - ethics of ineffectual drugs

modern impact 1/2

- ▶ FDA / drug approval
- ▶ ethics of the above, power of statistics v. expertise
  - ▶ ethics of placebo
  - ▶ ethics of harmful drugs
  - ▶ ethics of ineffectual drugs
- ▶ publication criterion

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo
  - ethics of harmful drugs
  - ethics of ineffectual drugs
- publication criterion
- in psychology (see Philips on transition from anecdote to data)

modern impact 1/2

- ▶ FDA / drug approval
- ▶ ethics of the above, power of statistics v. expertise
  - ▶ ethics of placebo
  - ▶ ethics of harmful drugs
  - ▶ ethics of ineffectual drugs
- ▶ publication criterion
- ▶ in psychology (see Philips on transition from anecdote to data)
- ▶ in economics

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo
  - ethics of harmful drugs
  - ethics of ineffectual drugs
- publication criterion
- in psychology (see Philips on transition from anecdote to data)
- in economics
- in engineering & reinforcement learning

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo
  - ethics of harmful drugs
  - ethics of ineffectual drugs
- publication criterion
- in psychology (see Philips on transition from anecdote to data)
- in economics
- in engineering & reinforcement learning
- reification of mathematical statistics

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
    - ethics of placebo
    - ethics of harmful drugs
    - ethics of ineffectual drugs
- publication criterion
- in psychology (see Philips on transition from anecdote to data)
- in economics
- in engineering & reinforcement learning
- reification of mathematical statistics
    - utility

modern impact 1/2

- FDA / drug approval
- ethics of the above, power of statistics v. expertise
  - ethics of placebo
  - ethics of harmful drugs
  - ethics of ineffectual drugs
- publication criterion
- in psychology (see Philips on transition from anecdote to data)
- in economics
- in engineering & reinforcement learning
- reification of mathematical statistics
  - utility
  - resources

modern impact 1/2

- ▶ FDA / drug approval
- ▶ ethics of the above, power of statistics v. expertise
    - ▶ ethics of placebo
    - ▶ ethics of harmful drugs
    - ▶ ethics of ineffectual drugs
- ▶ publication criterion
- ▶ in psychology (see Philips on transition from anecdote to data)
- ▶ in economics
- ▶ in engineering & reinforcement learning
- ▶ reification of mathematical statistics
    - ▶ utility
    - ▶ resources
    - ▶ professionalization + legitimization

- what even is "statistics"? at this point we've seen:

modern impact 2/2

- what even is "statistics"? at this point we've seen:
  - quantitative statecraft

modern impact 2/2

- what even is "statistics"? at this point we've seen:
  - quantitative statecraft
  - gather data, for others to explain or model

modern impact 2/2

- what even is "statistics"? at this point we've seen:
  - quantitative statecraft
  - gather data, for others to explain or model
  - social physics

modern impact 2/2

- what even is "statistics"? at this point we've seen:
    - quantitative statecraft
    - gather data, for others to explain or model
    - social physics
    - mathematical statistics

modern impact 2/2

- what even is "statistics"? at this point we've seen:
  - quantitative statecraft
  - gather data, for others to explain or model
  - social physics
  - mathematical statistics
  - science of uncertainty

modern impact 2/2

- what even is "statistics"? at this point we've seen:
    - quantitative statecraft
    - gather data, for others to explain or model
    - social physics
    - mathematical statistics
    - science of uncertainty
    - algorithmic truthmaking

themes for today

- early trading zone: industry, agriculture, and mathematics (itself localized to academia)

themes for today

- early trading zone: industry, agriculture, and mathematics (itself localized to academia)
- mathematical-objectivity as truthmaking culminates in algorithmic-objectivity as truth-making

themes for today

- early trading zone: industry, agriculture, and mathematics (itself localized to academia)
- mathematical-objectivity as truthmaking culminates in algorithmic-objectivity as truth-making
- art vs scale (NB: art requires artists, scale requires algorithms)

guiding questions every week:

- Scientific and mathematical development

guiding questions every week:

- ▶ Scientific and mathematical development
- ▶ Technologies and engineering (not so much in "part 1", more in parts 2/3 of class)

guiding questions every week:

- ▶ Scientific and mathematical development
- ▶ Technologies and engineering (not so much in "part 1", more in parts 2/3 of class)
- ▶ Driving forces: money, prestige, resources, Imperial competition... and "truth!"

guiding questions every week:

- Scientific and mathematical development
- Technologies and engineering (not so much in "part 1", more in parts 2/3 of class)
- Driving forces: money, prestige, resources, Imperial competition. . . and "truth!"
- how did new capabilities rearrange power? (who can now do what, from what, to whom?) (rights, harms, justice)

appendix

appendix

- 2020-01-21 : 1 of 14 intro to course

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- 2020-03-03 : 7 of 14 WWII, dawn of digital computation

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- 2020-03-10 : 8 of 14 birth and death of AI

appendix

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- 2020-03-10 : 8 of 14 birth and death of AI
- 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- 2020-03-31 : 10 of 14 data science, 1962-2017

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- 2020-03-10 : 8 of 14 birth and death of AI
- 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- 2020-03-31 : 10 of 14 data science, 1962-2017
- 2020-04-07 : 11 of 14 AI2.0

appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- ▶ 2020-03-31 : 10 of 14 data science, 1962-2017
- ▶ 2020-04-07 : 11 of 14 AI2.0
- ▶ 2020-04-14 : 12 of 14 ethics

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- 2020-03-10 : 8 of 14 birth and death of AI
- 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- 2020-03-31 : 10 of 14 data science, 1962-2017
- 2020-04-07 : 11 of 14 AI2.0
- 2020-04-14 : 12 of 14 ethics
- 2020-04-21 : 13 of 14 present problems: attention economy+VC=dumpsterfire

appendix

- 2020-01-21 : 1 of 14 intro to course
- 2020-01-28 : 2 of 14 setting the stakes
- 2020-02-04 : 3 of 14 risk and social physics
- 2020-02-11 : 4 of 14 statecraft and quantitative racism
- 2020-02-18 : 5 of 14 intelligence, causality, and policy
- 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- 2020-03-10 : 8 of 14 birth and death of AI
- 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- 2020-03-31 : 10 of 14 data science, 1962-2017
- 2020-04-07 : 11 of 14 AI2.0
- 2020-04-14 : 12 of 14 ethics
- 2020-04-21 : 13 of 14 present problems: attention economy+VC=dumpsterfire
- 2020-04-28 : 14 of 14 future solutions