data-ppf.github.io 2021-02-16

lecture 6 of 14: data gets real: mathematical baptism

chris wiggins + matt jones, Columbia

today: end of Part 1

today: end of Part 1

1. Gosset (b. 1876)

today: end of Part 1

1. Gosset (b. 1876)

- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)

today: end of Part 1

1. Gosset (b. 1876)

- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)

2. Fisher (b. 1890)

today: end of Part 1

1. Gosset (b. 1876)

- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)

2. Fisher (b. 1890)

- and eugenics

today: end of Part 1

1. Gosset (b. 1876)

- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)

2. Fisher (b. 1890)

- and eugenics

3. Neyman (b. 1894)

today: end of Part 1

1. Gosset (b. 1876)

- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)

2. Fisher (b. 1890)

- and eugenics

3. Neyman (b. 1894)

- and Bayes

today: end of Part 1

1. Gosset (b. 1876)
   - and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)
2. Fisher (b. 1890)
   - and eugenics
3. Neyman (b. 1894)
   - and Bayes
4. Lasting impact

today: end of Part 1

1. Gosset (b. 1876)
- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)
2. Fisher (b. 1890)
- and eugenics
3. Neyman (b. 1894)
- and Bayes
4. Lasting impact
- "mathematical statistics" finally eats "statistics" for good

today: end of Part 1

1. Gosset (b. 1876)

- and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)

2. Fisher (b. 1890)

- and eugenics

3. Neyman (b. 1894)

- and Bayes

4. Lasting impact

- "mathematical statistics" finally eats "statistics" for good
- the truth algorithm

today: end of Part 1

1. Gosset (b. 1876)
   - and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)
2. Fisher (b. 1890)
   - and eugenics
3. Neyman (b. 1894)
   - and Bayes
4. Lasting impact
   - "mathematical statistics" finally eats "statistics" for good
   - the truth algorithm
5. science -vs- engineering (and utilitarianism)

today: end of Part 1

1. Gosset (b. 1876)
   - and K. Pearson (b. 1857), "protégé and biographer" of Galton (b. 1822)
2. Fisher (b. 1890)
   - and eugenics
3. Neyman (b. 1894)
   - and Bayes
4. Lasting impact
   - "mathematical statistics" finally eats "statistics" for good
   - the truth algorithm
5. science -vs- engineering (and utilitarianism)
   - and WWII (start of Part 2)

but first. . . student reactions

but first. . . student reactions

```
4 "set in stone" (tests)
4 "over my head" (the mathy material)
2 sass/sassy
```

student reactions: statisticians are people too

- "I took great enjoyment in watching these two titans of statistics publicly beefing

student reactions: statisticians are people too

- "I took great enjoyment in watching these two titans of statistics publicly beefing
- "read like a mildly juicy episode of Keeping Up with the Kardashians or some other emotionally charged reality tv show.

student reactions: statisticians are people too

- "I took great enjoyment in watching these two titans of statistics publicly beefing
- "read like a mildly juicy episode of Keeping Up with the Kardashians or some other emotionally charged reality tv show.
- "These readings, more than any of those that preceded them, delved into the character of the statisticians in light of their findings. . . [Fisher] was religiously eugenic.

student reactions: statisticians are people too

- "I took great enjoyment in watching these two titans of statistics publicly beefing
- "read like a mildly juicy episode of Keeping Up with the Kardashians or some other emotionally charged reality tv show.
- "These readings, more than any of those that preceded them, delved into the character of the statisticians in light of their findings. . . [Fisher] was religiously eugenic.
- "it is noteworthy that just like today's big tech companies, the Guinness corporation sought to recruit the best talent from top universities

student reactions: hard/more stats

- "a little dry

student reactions: hard/more stats

- "a little dry
- "a lot of the statistical talk in this week's readings went right over my head

student reactions: hard/more stats

- "a little dry
- "a lot of the statistical talk in this week's readings went right over my head
- "those pieces were interesting, no doubt, but mathy

student reactions: hard/more stats

- "a little dry
- "a lot of the statistical talk in this week's readings went right over my head
- "those pieces were interesting, no doubt, but mathy
- "I couldn't make sense of the mathematical minutiae fueling this rift

student reactions: art+science (vs “objectivity”)

- "Fisher states that the choice of test statistic and null hypothesis was an art and not a mechanical process

student reactions: art+science (vs “objectivity”)

- "Fisher states that the choice of test statistic and null hypothesis was an art and not a mechanical process
- "we misunderstand statistics as objective truth and how dangerous or misleading it can be when we do

student reactions: art+science (vs “objectivity”)

- "Fisher states that the choice of test statistic and null hypothesis was an art and not a mechanical process
- "we misunderstand statistics as objective truth and how dangerous or misleading it can be when we do
- "my main takeaway from this week’s readings: the way we have been taught statistics cannot be viewed as objective.

student reactions: art+science (vs “objectivity”)

- "Fisher states that the choice of test statistic and null hypothesis was an art and not a mechanical process
- "we misunderstand statistics as objective truth and how dangerous or misleading it can be when we do
- "my main takeaway from this week’s readings: the way we have been taught statistics cannot be viewed as objective.
- "I had always taken measures of error and significant tests (such as t-tests, standard deviation, etc) as true and objective

student reactions: how to relate to present day, other classes?

these, then, were vigorous controversies, and they have not ended. . . so it is remarkable that all of these unresolved controversial issues, conceptual ambiguities, and personal insults have been more or less completely suppressed from the textbooks" (EoC p. 105).

recurring theme: capabilities and consequences

I want to echo $student's post, as I think he asks a question that we as a class should try to answer: "Can it be considered a victory that the field of statistics embraced mathematical rigor if the applications of such an embrace might be used to discriminate with an unmatched precision?" There have been many "breakthroughs" in science and mathematics stemming from unethical means or immoral motivations. To what extent do we feel comfortable acknowledging the circumstances that have led to such breakthroughs and to what extent should we teach these circumstances to prevent them from recurring in future?

1. Gossett & Guinness (1899)

- 1886 IPO

1. Gossett & Guinness (1899)

- 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.

1. Gossett & Guinness (1899)

- 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.
- 1.5T 2021 USD (adjusting for inflation)

1. Gossett & Guinness (1899)

- 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.
- 1.5T 2021 USD (adjusting for inflation)
- approx value of GOOG

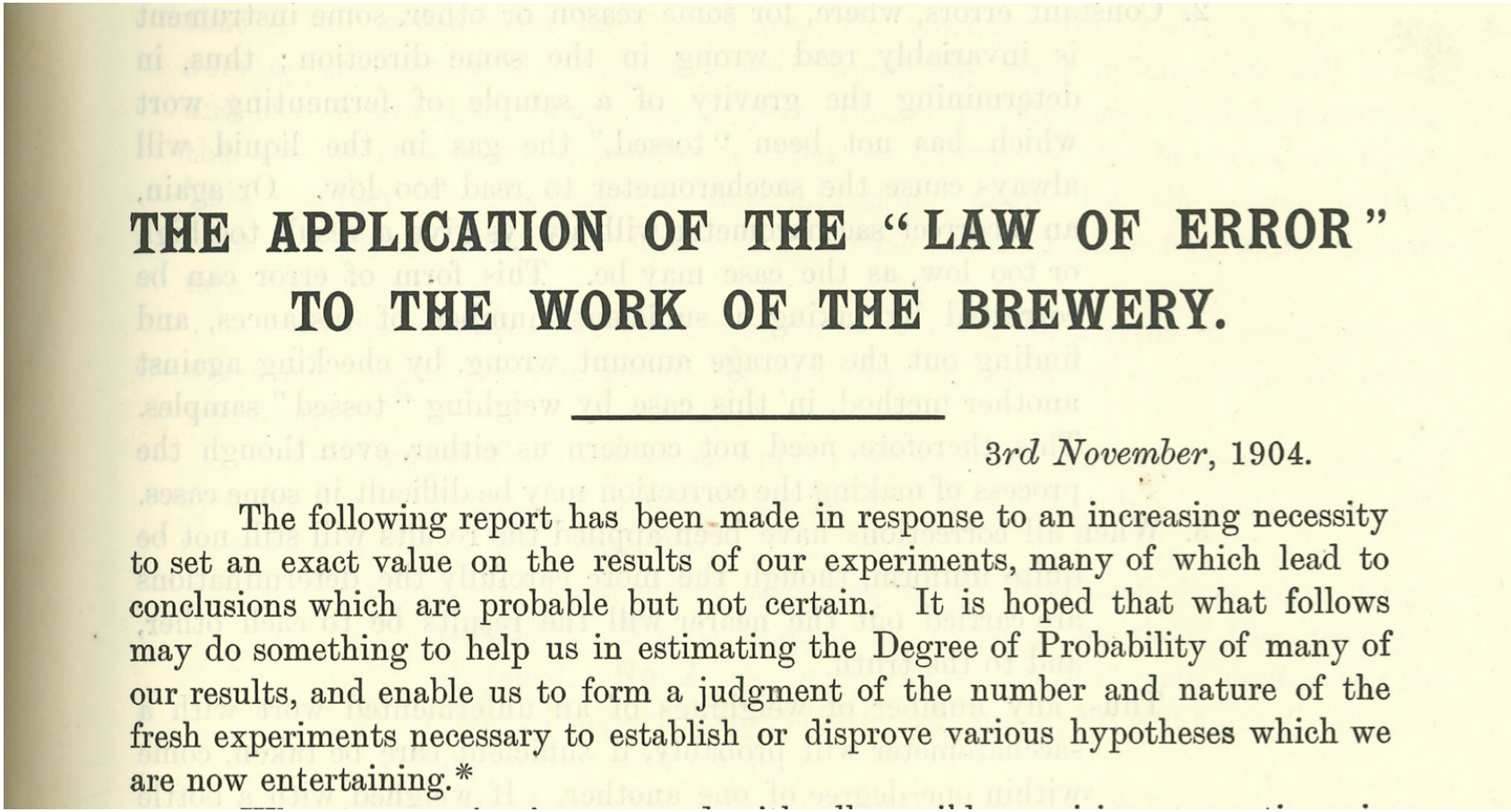
1. Gossett & Guinness (1899)

- 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.
- 1.5T 2021 USD (adjusting for inflation)
- approx value of GOOG
- Guinness recruits the "elite" students, treats them well

1. Gossett & Guinness (1899)

- 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.
- 1.5T 2021 USD (adjusting for inflation)
- approx value of GOOG
- Guinness recruits the "elite" students, treats them well
- The brewers "had become keen geneticists and breeders since the rediscovery in 1900 of Mendel's paper on hybridization"

Statistical tools & agricultural experiment



THE APPLICATION OF THE "LAW OF ERROR" TO THE WORK OF THE BREWERY.

3rd November, 1904.

The following report has been made in response to an increasing necessity to set an exact value on the results of our experiments, many of which lead to conclusions which are probable but not certain. It is hoped that what follows may do something to help us in estimating the Degree of Probability of many of our results, and enable us to form a judgment of the number and nature of the fresh experiments necessary to establish or disprove various hypotheses which we are now entertaining.*

Figure 1: Gosset

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

1905: Gossett and K. Pearson

- " in July 1905 during Gosset's summer holiday . . . when he cycled about 20 miles from his parents' home to Pearson's summer home, . . . [learning] "nearly all the methods then in use."

1905: Gossett and K. Pearson

- " in July 1905 during Gosset's summer holiday . . . when he cycled about 20 miles from his parents' home to Pearson's summer home, . . . [learning] "nearly all the methods then in use."
- Pearson: "correlation" today is "Pearson correlation coefficient"

1905: Gossett and K. Pearson

- " in July 1905 during Gosset's summer holiday . . . when he cycled about 20 miles from his parents' home to Pearson's summer home, . . . [learning] "nearly all the methods then in use."
- Pearson: "correlation" today is "Pearson correlation coefficient"
- "the first holder of this chair — the Galton Chair of Eugenics, later the Galton Chair of Genetics"

1905: Gossett and K. Pearson

- " in July 1905 during Gosset's summer holiday . . . when he cycled about 20 miles from his parents' home to Pearson's summer home, . . . [learning] "nearly all the methods then in use."
- Pearson: "correlation" today is "Pearson correlation coefficient"
- "the first holder of this chair — the Galton Chair of Eugenics, later the Galton Chair of Genetics"
- Student's intellectual advance: "t-test" of 1908

1905: Gossett and K. Pearson

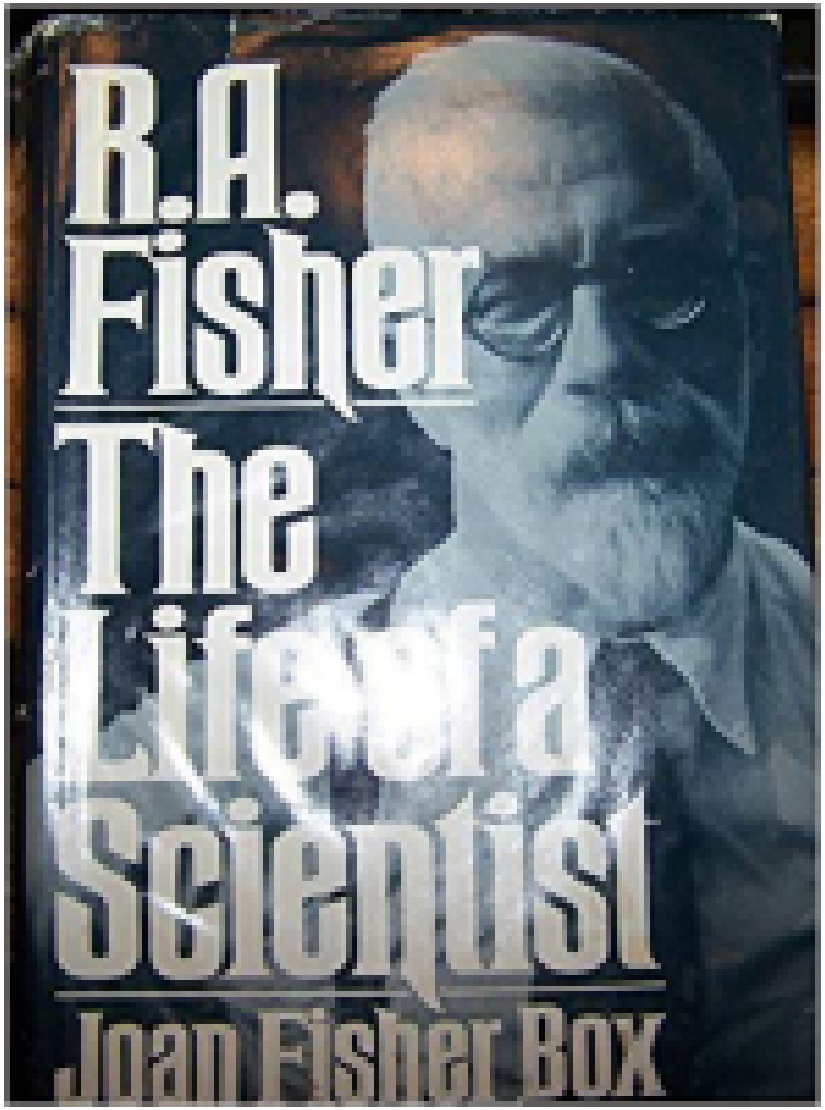
- " in July 1905 during Gosset's summer holiday . . . when he cycled about 20 miles from his parents' home to Pearson's summer home, . . . [learning] "nearly all the methods then in use."
- Pearson: "correlation" today is "Pearson correlation coefficient"
- "the first holder of this chair — the Galton Chair of Eugenics, later the Galton Chair of Genetics"
- Student's intellectual advance: "t-test" of 1908
- markedly different epistemic goals, early trading zone: industry, agriculture, and mathematics (itself localized to academia)

1912: Gossett gets letter from young (22 y/o) Fisher

> *as Gosset wrote of it to Pearson on 12 September 1912 (Pearson, 1968): " Stratton, the tutor, made him send me and with some exertion I mastered it, I couldn't understand his stuff and wrote and said I was going to study it when I had time. I actually took it up to the lakes with me—and lost it! It's so nice and mathematical that it might appeal to some people.*

Wikip: From 1912 to 1934 Gosset and Fisher would exchange more than 150 letters. In 1924, Gosset wrote in a letter to Fisher, "I am sending you a copy of Student's Tables as you are the only man that's ever likely to use them!"

2. Fisher (b. 1890)



2. Fisher (b. 1890)

- RAF: Rothamsted->UCL (eugenics)->Cambridge->Australia

2. Fisher (b. 1890)

- RAF: Rothamsted->UCL (eugenics)->Cambridge->Australia
- Fisher formed Cambridge Eugenics Society as sophomore: *"The widespread, fruitful, and successful races of the future belong to the dominant nations of to-day"*

Fisher: big ag @ Rothamsted Experimental Station 1/2

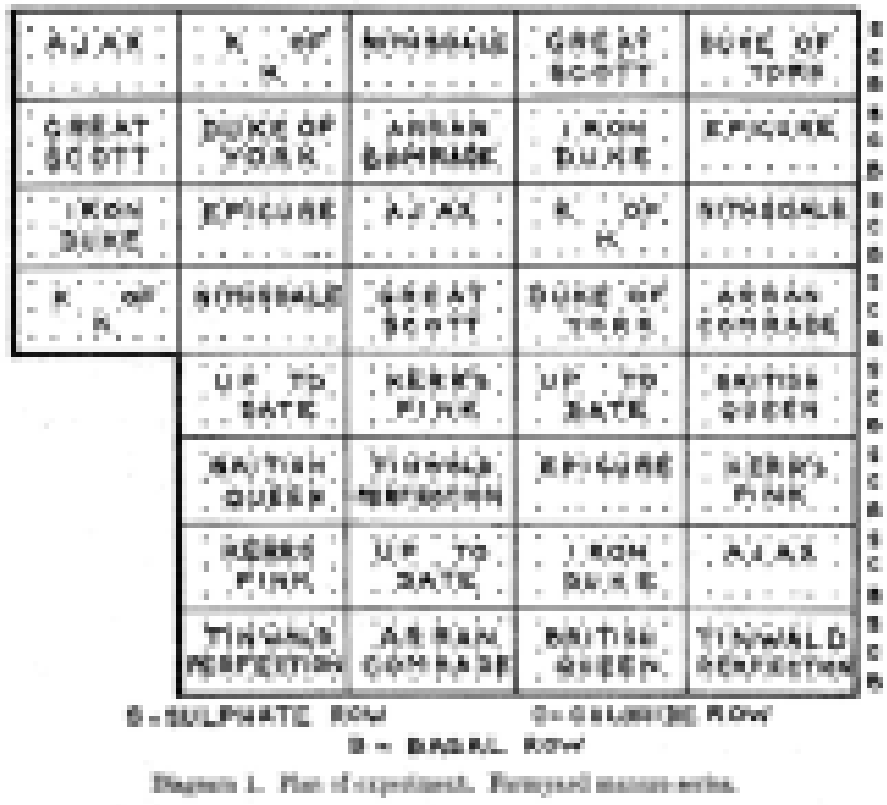


Figure 2: Fisher-Mackenzie crop experimental design

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 3: "millionaire calculator"

"Millionaire calculator... was equivalent to six months of his own salary" 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 3: "millionaire calculator"

"Millionaire calculator... was equivalent to six months of his own salary" 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 3: "millionaire calculator"

"Millionaire calculator... was equivalent to six months of his own salary" purchased for Fisher ££££$$$$ tech of the day!

Fisher: not a mathphobe

- "He entered Harrow School age 14 and won the school's Neeld Medal in mathematics.

(from Wikip)

Fisher: not a mathphobe

- "He entered Harrow School age 14 and won the school's Neeld Medal in mathematics.
- In 1909, he won a scholarship to study Mathematics at Gonville and Caius College, Cambridge.

(from Wikip)

Fisher: not a mathphobe

- "He entered Harrow School age 14 and won the school's Neeld Medal in mathematics.
- In 1909, he won a scholarship to study Mathematics at Gonville and Caius College, Cambridge.
- In 1912, he gained a First in Mathematics."

(from Wikip)

Fisher: experiment requires math, 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 all results which fail to reach this standard."

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 all results which fail to reach this standard."
- NB: "significance testing" not "hypothesis testing"

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."

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: insistence on randomization

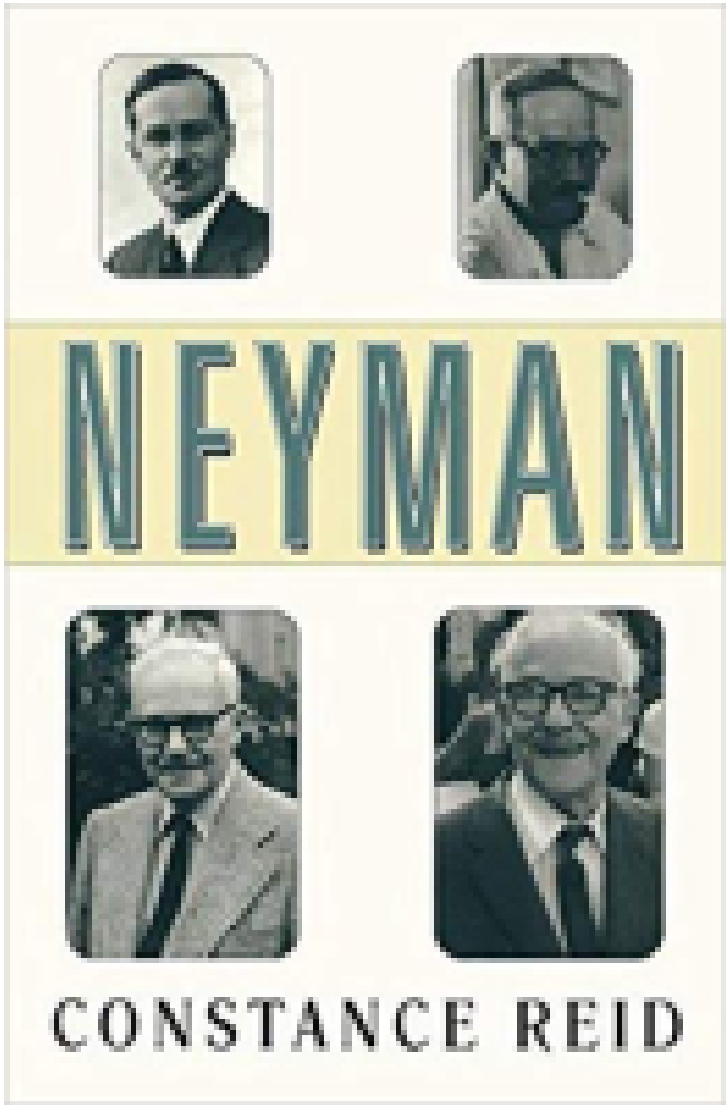
- Randomization necessary so that "the test of significance may be guaranteed against corruption by the causes of disturbance which have not been eliminated."
- Randomization "relieves the experimenter from the anxiety of considering and estimating the magnitude of the innumerable causes by which his data may be disturbed."
- Q: how does this relate to causality?

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."

3. Neyman



3. Neyman

- interwar Poland

3. Neyman

- interwar Poland
- US 1938

3. Neyman

- interwar Poland
- US 1938
- Berkeley statistical lab

3. Neyman

- interwar Poland
- US 1938
- Berkeley statistical lab
- Berkeley statistical department

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 $\neq$ 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: science $\neq$ 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: 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)

3.a Math, "type 1/type 2 errors"

mlstory.org/decision.html

conditional densities are

$$p(X \mid H_0 \text{ is true}) = \mathcal{N}(0, 1)$$
$$p(X \mid H_1 \text{ is true}) = \mathcal{N}(s, 1)\,.$$

When s has large magnitude, it would be obvious whether H_0 or H_1 were true. For example, suppose $s = 10$ and we observed $X = 11$. Under H_0, the probability that the observation greater than 10 is on the order of 10^{-23}, and hence we'd likely think we're in alternative H_1 . However, if s were very close to zero, distinguishing between the two alternatives is rather challenging. We can think of a small signal s that we're trying to detect as a *needle in a haystack*.

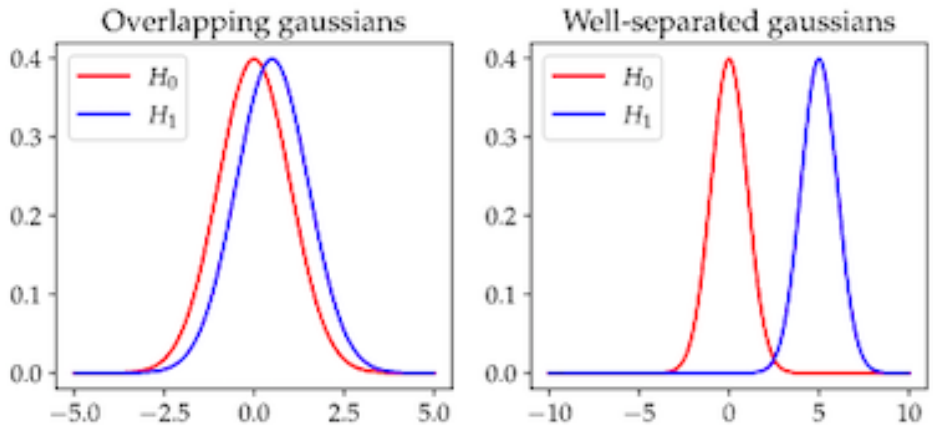


Figure 4: new book: mlstory.org

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'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'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

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

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.

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.
- but the worst insult is: you're Bayesian

3.b Bayes: everyone hates him

Figure 5: Reverend Bayes. We will need him when we go to war (next week)

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

- consider university w/10,000 students

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

- consider university w/10,000 students
- "perfect" test assesses 1% have disease

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

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

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

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

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

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

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

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

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

	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

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

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

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

	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)

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 6: Stigler on Bayes on $p(M|t)$ and 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 6: Stigler on Bayes on $p(M|t)$ and God

everyone hates Bayes

- Fisher willing as long as prior itself is known via sampling

everyone hates Bayes

- Fisher willing as long as prior itself is known via sampling
- else: “subjective. . . taste”

hypotheses: what are they *good* for?

- see EoC example:

hypotheses: what are they *good* for?

- 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.

hypotheses: what are they *good* for?

- 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.
- $\mu(a|d) = \sum_h R(a,h)p(h|d) = \frac{1}{p(d)} \sum_h R(a,h)p(d|h)p(h)$

4. Lasting impact

4a. Placement of stats

- what even is “statistics”? at this point we’ve seen:

4a. Placement of stats

- what even is “statistics”? at this point we’ve seen:
  - quantitative statecraft

4a. Placement of stats

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

4a. Placement of stats

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

4a. Placement of stats

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

4a. Placement of stats

- 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

4a. Placement of stats

- 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

4a. Placement of stats

The growing need, demand, and opportunity have confronted the educational system of the country with a series of problems regarding the teaching of statistics. Should statistics be taught in the department of agriculture, anthropology, astronomy, biology, business, economics, education, engineering, medicine, physics, political science, psychology, or sociology, or in all these departments? Should its teaching be entrusted to the department of mathematics, or a separate department of statistics, and in either of these cases should other departments be prohibited from offering duplicating courses in statistics, as they are often inclined to do?

Figure 7: Hoteling, 1945

4a. Placement of stats

2. *Physical, Mathematical, and Engineering Sciences.* For this report (a) *physical sciences* are those sciences concerned primarily with the understanding of the natural phenomena associated with nonliving things; (b) *mathematical sciences* are those sciences which employ logical reasoning with the aid of symbols and which are concerned with the development of methods of operations employing such symbols, including mathematics, pure and applied; astronomy, theoretical mechanics, statistics, logistic research, and computer research exclusive of engineering; (c) *engineering sciences* are those sciences which are concerned with studies directed toward making specific scientific principles usable in engineering practice.

Figure 8: NSF, 1952

4b. The Truth Algorithm

- FDA / drug approval

4b. The Truth Algorithm

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

4b. The Truth Algorithm

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

4b. The Truth Algorithm

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

4b. The Truth Algorithm

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

4b. The Truth Algorithm

- 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

4b. The Truth Algorithm

- 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)

4b. The Truth Algorithm

- 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

4b. The Truth Algorithm

- 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

4b. The Truth Algorithm

- 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

4b. The Truth Algorithm

- 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

4b. The Truth Algorithm

- 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

4b. The Truth Algorithm

- 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

FDA/drug approval: a sidebar 1/3

- "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.

FDA/drug approval: a sidebar 1/3

- "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.

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

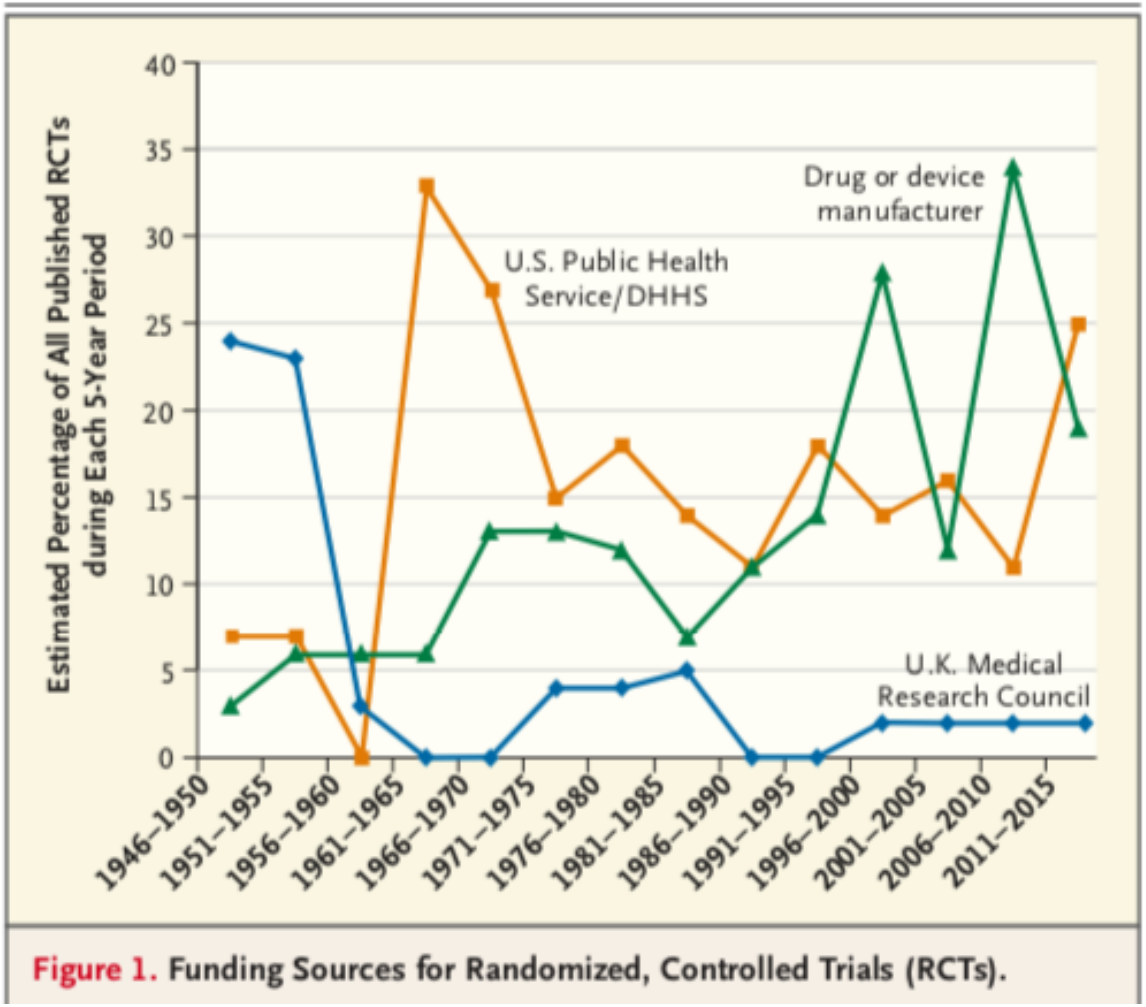


Figure 9: percent of RCT from US gov, UK gov, industry

4.b. Truth algorithm: pushback

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND *P*-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative Science

March 7, 2016

Figure 10: ASA statement, March 2016

5. Utilitarianism

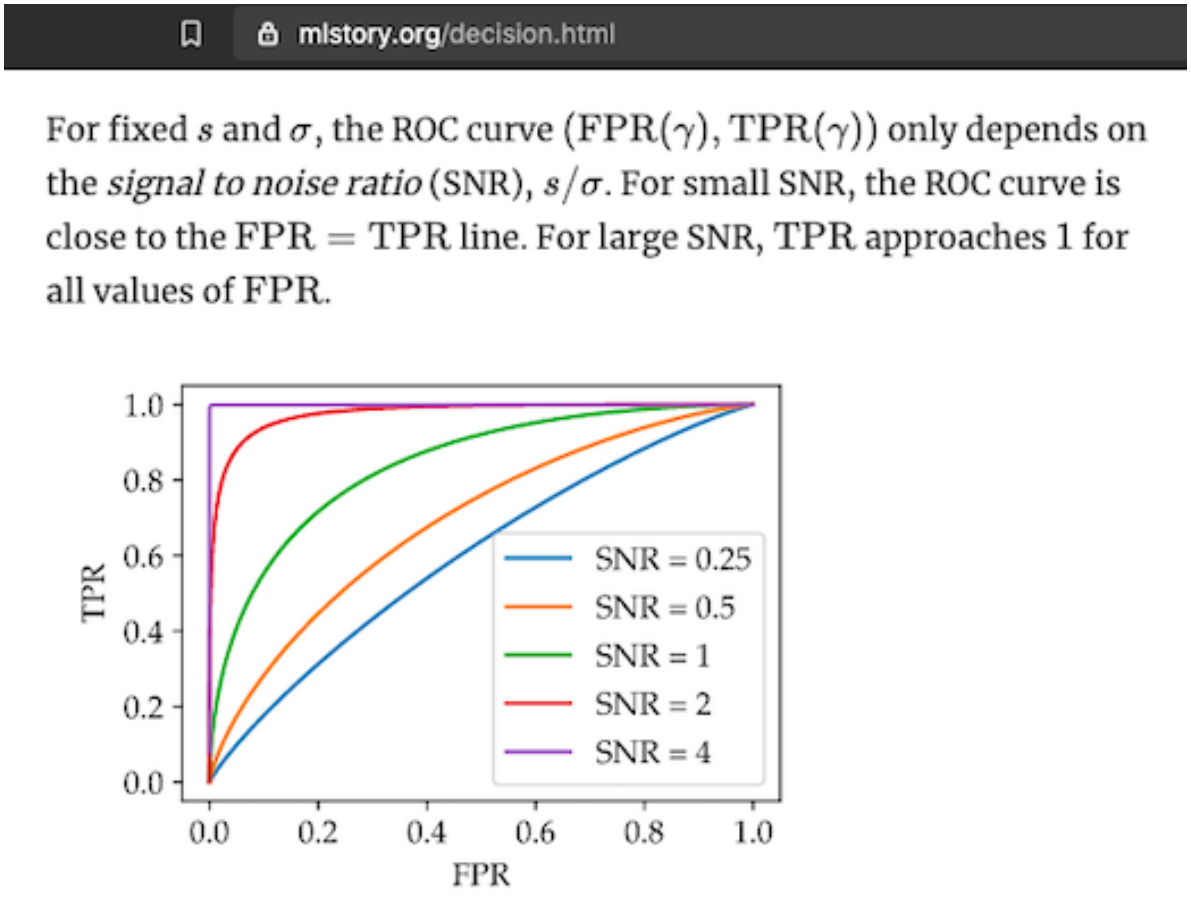
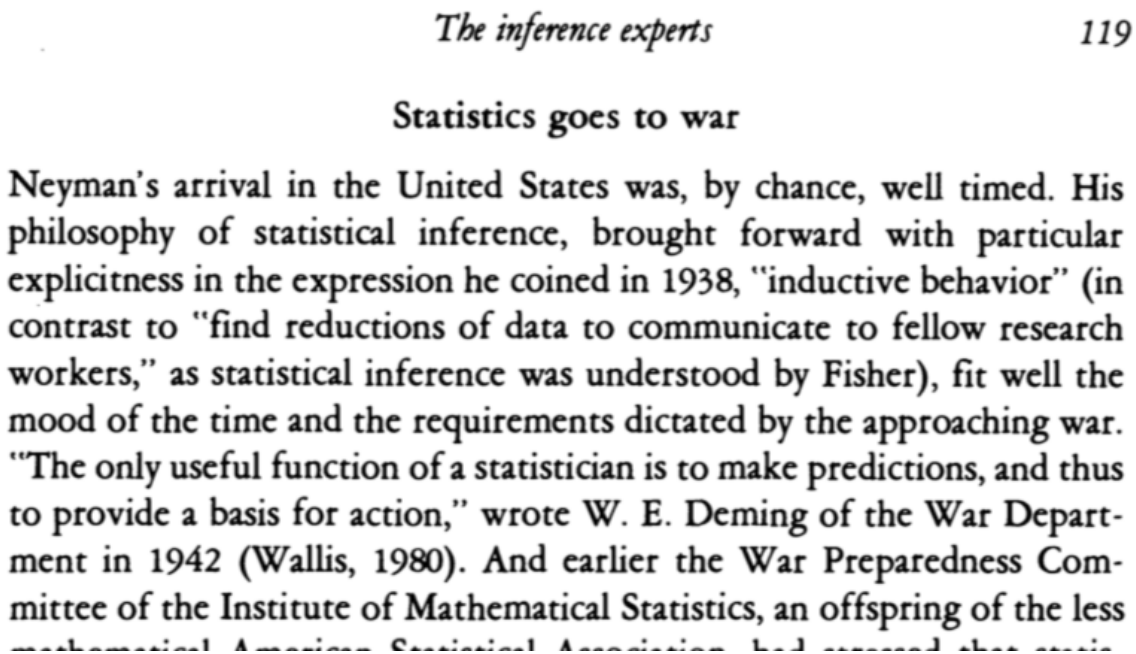


Figure 11: new book: mlstory.org

5. Utilitarianism and WWII



The inference experts 119

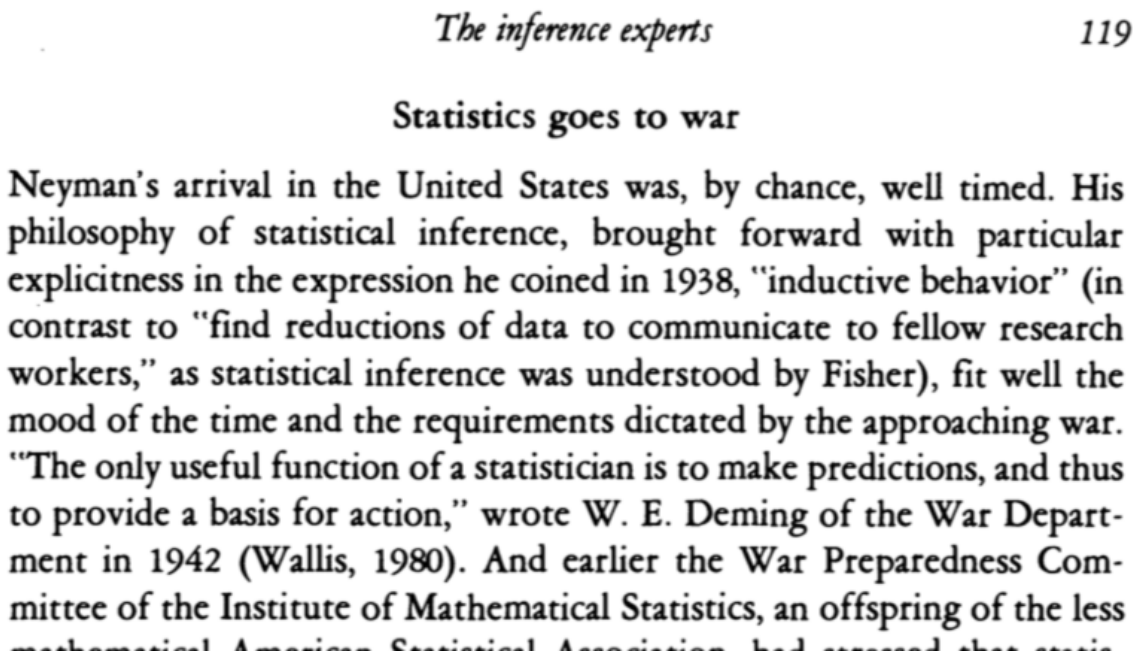
Statistics goes to war

Neyman's arrival in the United States was, by chance, well timed. His philosophy of statistical inference, brought forward with particular explicitness in the expression he coined in 1938, "inductive behavior" (in contrast to "find reductions of data to communicate to fellow research workers," as statistical inference was understood by Fisher), fit well the mood of the time and the requirements dictated by the approaching war. "The only useful function of a statistician is to make predictions, and thus to provide a basis for action," wrote W. E. Deming of the War Department in 1942 (Wallis, 1980). And earlier the War Preparedness Committee of the Institute of Mathematical Statistics, an offspring of the less

Figure 12: Deming, quoted by Gerd, on stats in war

- next week: data and War:

5. Utilitarianism and WWII



The inference experts 119

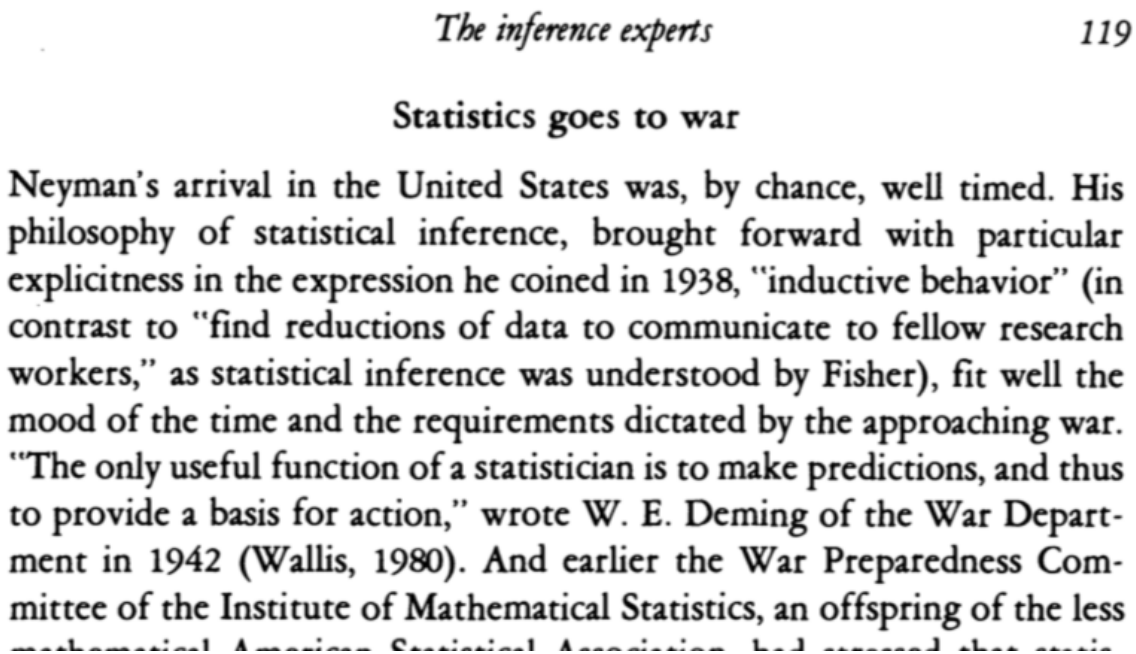
Statistics goes to war

Neyman's arrival in the United States was, by chance, well timed. His philosophy of statistical inference, brought forward with particular explicitness in the expression he coined in 1938, "inductive behavior" (in contrast to "find reductions of data to communicate to fellow research workers," as statistical inference was understood by Fisher), fit well the mood of the time and the requirements dictated by the approaching war. "The only useful function of a statistician is to make predictions, and thus to provide a basis for action," wrote W. E. Deming of the War Department in 1942 (Wallis, 1980). And earlier the War Preparedness Committee of the Institute of Mathematical Statistics, an offspring of the less

Figure 12: Deming, quoted by Gerd, on stats in war

- next week: data and War:
  - the birth of digital computation

5. Utilitarianism and WWII



The inference experts 119

Statistics goes to war

Neyman's arrival in the United States was, by chance, well timed. His philosophy of statistical inference, brought forward with particular explicitness in the expression he coined in 1938, "inductive behavior" (in contrast to "find reductions of data to communicate to fellow research workers," as statistical inference was understood by Fisher), fit well the mood of the time and the requirements dictated by the approaching war. "The only useful function of a statistician is to make predictions, and thus to provide a basis for action," wrote W. E. Deming of the War Department in 1942 (Wallis, 1980). And earlier the War Preparedness Committee of the Institute of Mathematical Statistics, an offspring of the less

Figure 12: Deming, quoted by Gerd, on stats in war

- next week: data and War:
  - the birth of digital computation
  - engineering and labor, including Ghost Work

Today's themes, recap

Appendix

Appendix

- 2021-01-12: intro to course

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: data science, 1962-2017

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: data science, 1962-2017
- 2021-03-30: AI2.0

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: data science, 1962-2017
- 2021-03-30: AI2.0
- 2021-04-06: ethics

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: data science, 1962-2017
- 2021-03-30: AI2.0
- 2021-04-06: ethics
- 2021-04-13: present problems: attention economy+VC=dumpsterfire

Appendix

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: data science, 1962-2017
- 2021-03-30: AI2.0
- 2021-04-06: ethics
- 2021-04-13: present problems: attention economy+VC=dumpsterfire
- 2021-04-15: future solutions