Ph458/213 Evidence and Policy
Lecture 2

John Worrall
LAK 3.02
Office Hours:
Monday 13.30-14.30
Thursday12.30-13.30

Surrogate outcomes

- Real story these deaths were iatrogenic: caused by antiarrythmic drugs.
- The story of ventricular ectopic beats
- You DON'T want to have your VEBs suppressed
- You DO want to avoid cardiac arrest.
- Question 5: Is there evidence that there is a connection between the intervention and the outcome I really care about?
- (Easy mistake 5: settle for connections between the evidence and *surrogate* outcomes –
- Invariably because you feel it's plausible that there is a causal link between the surrogate and the real outcome
- But you need EVIDENCE for that
- AND you would like some evidence about a possible independent effect)
- Cp statins and reduction in cholesterol.

External Validity

- STAR project in Tennessee (1985-), RCT study ("gold standard of evidence")
- Result: reducing class sizes in Tennessee High Schools produced notable increases in educational performance
- (minority and inner city children benefitting 2 or 3 times as much as their white and nonurban peers)
- Formed an important basis for the decision by the California State Authorities in the 1990s to agree to spend $1bn p.a. (rising to $1.6bn) on reducing class sizes
- Result: no effect
- No suggestion that STAR produced the wrong result in Tennessee
- Question 6: Does the evidence generalise?
- (Easy mistake 6 assuming that because there is evidence that a policy worked 'there', it will also work 'here')

Evidence in Science

- Let’s begin by thinking about evidence in “hard science”
- Even in physics we can’t expect to establish theories conclusively (show that they are certainly true) on the basis of evidence.
- The reason for this is logical:
- Theories *go beyond* any possible evidence
- (a) in being universal
- [whereas the evidence we can collect is inevitably finite]
- (b) in making claims about what is going on ‘beyond’ or ‘beneath’ the data recorded in the evidence – to produce those data.

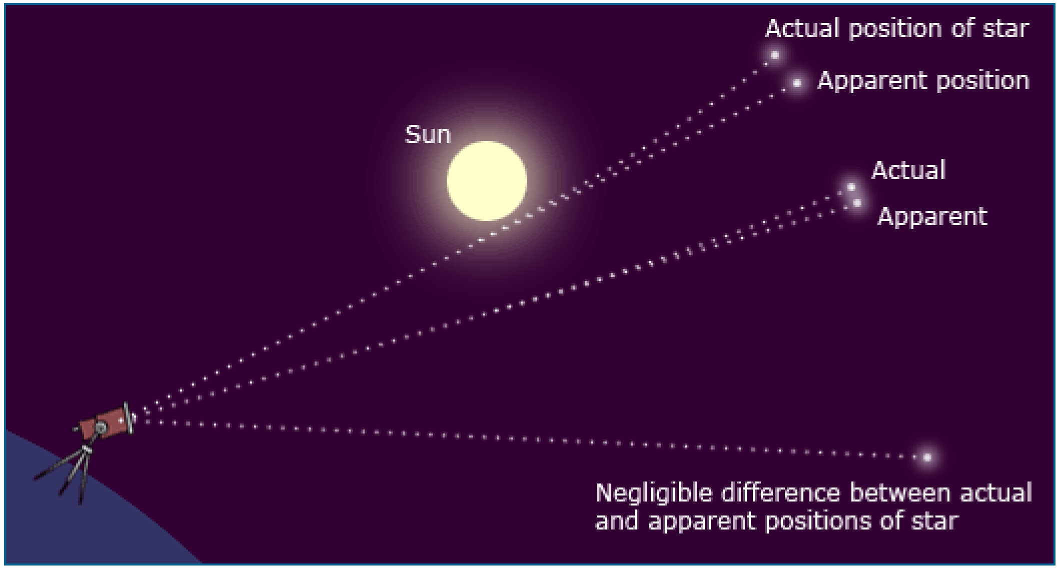
Evidence in Science

- But also illustrated historically
- Pope:
- *Nature and Nature’s Laws lay hid in night*
- *God said ‘Let Newton be!’ and all was light.*
- But Newton’s theory is not true (or so the evidence now says)

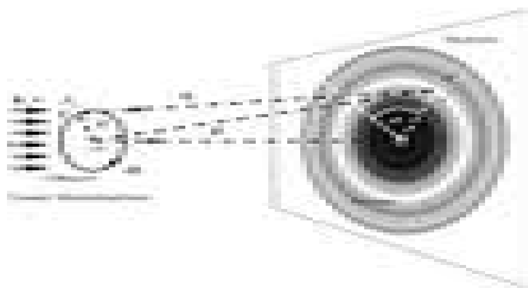
Evidence in Science

- Nonetheless we do sometimes think we have good evidence that a theory is at least approximately true.
- How?
- Via confirmation through *tests*
- A couple of examples

GTR and stellar positions



Wave theory of light and the ‘white spot’

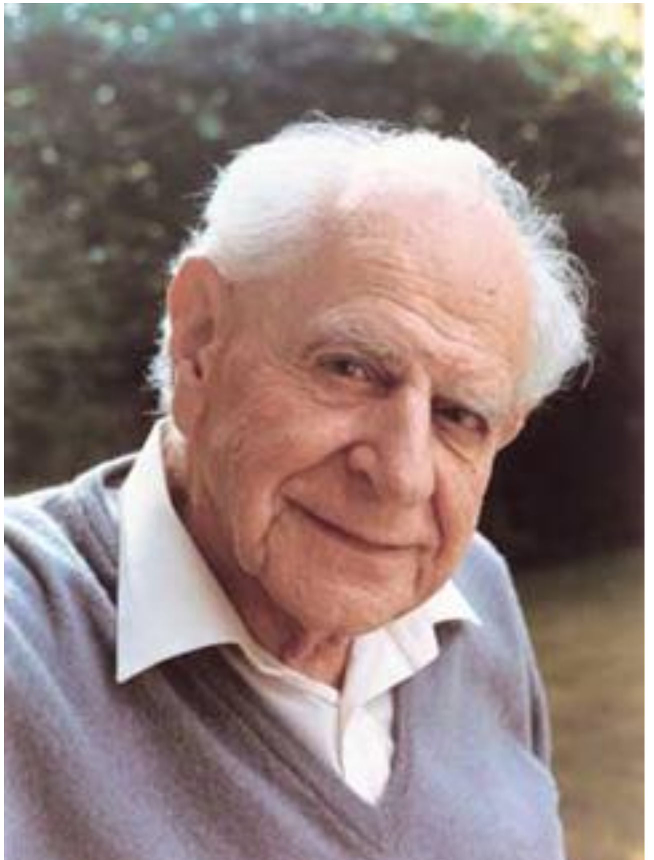


Evidence in Science

- These are cases where NOT ONLY
- 1. The theory entails the evidence
- BUT ALSO (and crucially)
- 2. The evidence is ***surprising*** - we would judge the probability of the evidence being as it is very low, *or* the probability of a theory getting that evidence correct would be very low, unless the theory at issue were true.
- Cases where condition 2 ***fails*** to hold:
- (a) Gypsy Rose Lee: “you will have a pleasant surprise tomorrow”
- (b) The “Gosse Dodge”

Evidence in Science

- This second – “otherwise surprising” – requirement is a part of every serious account of confirmation in science
- E.g. Popper
- Talks about ‘severe tests’



Evidence in Science

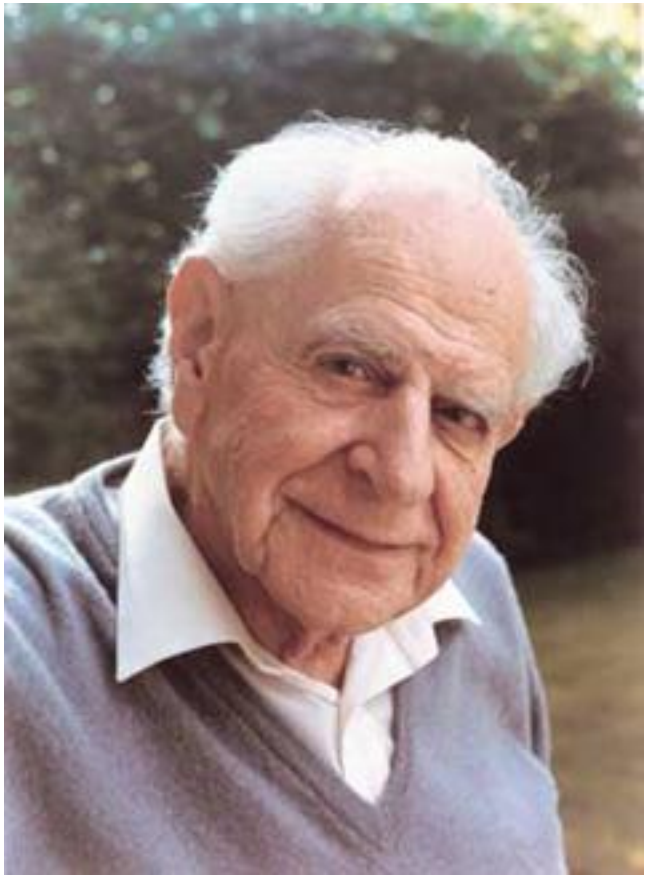
- And the Bayesian approach:
- The amount of confirmation lent by e to h
- Is inversely proportional to p(e)

Evidence in Science- Positive evidence

- So good/strong evidence FOR a theory – should elicit the ‘wow factor’: the theory must have something going for it if it can get a fact like that right!

Evidence in Science – negative evidence

- How about evidence AGAINST a theory?
- Popper's view



Evidence in Science – negative evidence

- Imre Lakatos:
- Things are more complicated
- Scientists hold on to their theories –
- despite ‘refutations’ -
- two types of outcome:
- (a) progressive (e.g. Neptune – wow factor returns!)
- (b) degenerative (e.g. classical physics, Gosse, the ‘French Paradox’ (??))

The ‘French Paradox’

- Hypothesis: Eating excess saturated fat in the diet raises LDL (Low Density Lipoprotein) levels . The LDL (aka "bad cholesterol") then causes thickening and narrowing in the arteries.
- The 'French Paradox‘: Over 15% of the calories in the average Frenchman's diet are supplied by saturated fats; this is a greater percentage than in Austria, Finland, Iceland, Switzerland .. all of whom have a higher rate of deaths from heart disease.
- Reaction: there must be something else about the French diet that compensates
- Leading to: any number of studies (e.g. into the effects of high garlic consumption) all of which have so far proved negative.

Testing *Statistical* Theories

- Ok so one way or the other tests play the crucial role in providing evidence for or against basic theories in physics
- But let’s come down to earth
- Relatively few practical policy decisions are based on such theories

Testing *Statistical* Theories

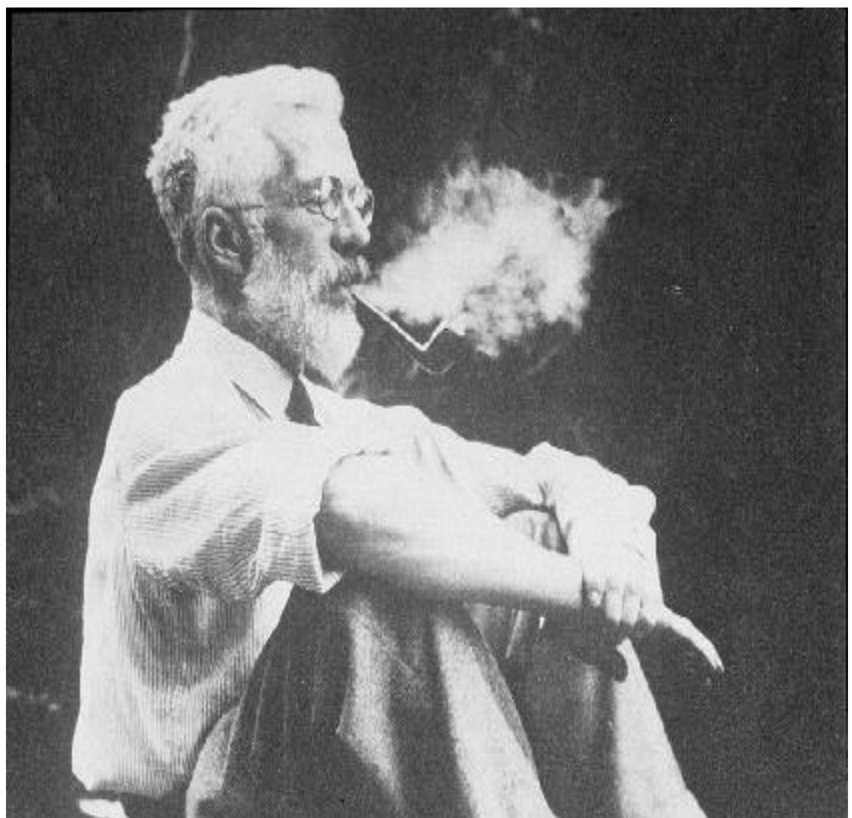
- Most theories of everyday concern and most relevant to policy issues are probabilistic/statistical:
- 1. You have a better chance of avoiding heart attacks if you take statins
- 2. You are more likely to develop lung cancer if you smoke cigarettes.
- 3. The probability of major flooding in London has been reduced by 25% by the Thames Barrier
- ....

Deterministic vs statistical theories

- There is clearly an important distinction
- Newton's theory is *deterministic*: given a complete specification S of the system at time t, the theory entails that the system *must* be in state S' at later time t'.
- So, e.g., a system consisting of a single body of mass 2 kgs subjected to a constant total force of 6 Newtons, will accelerate constantly at 3 m per sec per sec
- Clearly not true of the statistical theories I listed or in general
- Clear how we test deterministic theories; and – at least in outline – how we get evidence for or against them by testing
- Can we also get evidence for or against statistical theories by *testing* them?

Testing statistical theories

- This man, R.A. Fisher, insisted that we can
- And indeed that testing – 'significance testing' – was the very heart of statistics and statistical inference



Testing statistical theories

- An examination of the logic of Fisher’s – enormously influential – views will form an important part of this course.
- But first we should take a step backwards to think a bit about *probability*
- We need first to get straight on what probabilities *are*, before examining Fisher’s views on how probabilistic claims can be tested.

Probability Theory

- Game: drawing a card "at random" from a well-shuffled pack
- Outcome space/basic events
- Other events are then characterised in terms of basic events:
- E.g. event of drawing a club
- Or drawing a red card ..
- Probabilities of non-basic events
- Addition law?
- Real addition law:
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- [IF A and B are mutually exclusive $P(A \cup B) = P(A) + P(B)$]

Probability Theory

- $P(A \cap B)$?? (equivalently $P(A\&B)$)
- probability of drawing a card that is *both* a club *and* at the same time at least as high in value as a Jack?
- Only 4 cards satisfy both conditions so P (Club & ≥Jack) = 4/52
- How does this relate to P(Club) and P(≥Jack)?
- Simpler case: tossing a fair coin twice
- $P(H_1) = P(H_2) = ½$
- $P(H_1 \& H_2) = ¼$
- So(??) $P(A\&B) = P(A).P(B)$??
- Gives right result also with P (Club & ≥Jack)

Probability Theory

- BUT
- P(Heart & Red) ?
- Intuitively problem is that there is a connection between ‘Heart’ and ‘Red’
- Cp 1st and 2nd tosses of coin

Probability Theory

- Probabilistic dependence/independence
- ***Correct Law for Joint Probabilities***: for any two events A and B, P(A&B) = P(A).P(B/A)
- Special case: IF A and B are *independent*, then P(A&B) = P(A).P(B)
- Gives correct answers in all cases – e.g. P(Heart & Red)
- Since A & B is the same event as B & A we should also have
- for any two events A and B, P(A&B) = P(B).P(A/B)
- And that must imply P(A/B) = P(A) iff P(B/A) = P(B)
- Finally correct law yields a ‘definition’ of conditional probability
- P(B&A) = P(B).P(A/B) and so
- P(A/B) = P(B&A)/ P(B) = P(A&B)/ P(B) ***IF*** P(B) $\neq$ 0
- [For Probability axioms and some theorems – see further reading.]

“Inverse Probability”

- Back to coin tossing:
- Given: coin is fair
- Trial: toss it 4 times
- Statistic: r= number of heads out of 4
- Given that the coin is fair we can readily calculate the probabilities (“direct” probabilities) for all possible values of r
- P(4 heads) = P(0 heads) = $(½)^4$ = 1/16
- P(3 heads) = P(1 head) = 4/16
- P(2 heads) = 6/16

“Inverse Probability”

- Now suppose we have another coin which we are told is biassed
- In fact P(H) = ¾
- Same trial, same statistic
- Again easy to calculate the direct probabilities
- P(4 heads) = $(¾)^4$ = 81/256
- P(3heads) = $4.(1/4).(3/4)^3$
- P(2heads) = $6.(1/4)^2.(3/4)^2$
- P(1 head) = $4. (1/4)^3.(3/4)$
- P(0 heads) = $(1/4)^4$

“Inverse Probability”

- But ***now*** suppose we have both coins in a box, that they are physically indistinguishable, and that we draw one at random and then toss it 4 times
- Suppose the outcome is r = 1
- ***What’s the probability that it was the fair(biassed) coin that we tossed to produce that result?***
- Bayes’ Theorem allows us to answer.
- Bayes’ Theorem: P(A/B) = P(B/A).P(A)/ P(B).

“Inverse Probability”

- Bayes’ Theorem: P(A/B) = P(B/A).P(A)/ P(B).
- Let A be the event that it was the fair coin that we tossed
- (So ¬A is the event of having tossed the biased coin)
- And let B be the observed event: r =1
- We know P(B/A) = ¼
- And, given that the coin was chosen from the box ‘at random’ [??] P(A) = ½
- So we need one more probability to apply the theorem – viz
- P(B)
- What was the probability of getting one head out of 4 *whichever* coin was tossed?

"Inverse Probability"

- It's a general result ('Theorem on Total Probability') that
- P(B) = P(A).P(B/A) + P(¬A). P(B/¬A)
- So here
- P(B) = ½.P(B/A) + ½.P(B/¬A)
- (Why this is intuitively correct)
- So P(B) = ½ . ¼ + ½ . 4. $(1/4)^3$.(3/4) = 38/256
- So, remember, Bayes' Theorem is
- P(A/B) = P(B/A).P(A)/ P(B)
- Plugging in these values we get
- P(A/B) = ¼. ½ /(36/256) = 256/304
- So we have a 'prior' of ½ and a 'posterior' of over 2/3
- So probability has gone up that it is the fair coin even though an event occurred that was relatively unlikely to occur if it was the fair coin.
- (Why this is intuitively correct.)

The Bayesian approach to confirmation

- Rewriting Bayes' theorem to apply to the case of general interest where we have some hypothesis h and some evidence e, we have
- $P(h/e) = (P(e/h).P(h))/P(e)$
- $P(e/h)$ is the "likelihood" of the evidence in the light of h
- $P(h)$ is the "prior probability" of h
- $P(e)$ is the "prior probability" of e
- ***Fundamental Bayesian Principle:*** e confirms [or gives evidence in favour of] h if and only if $P(h/e) > P(e)$
- ***Fundamental Bayesian Principle:*** e confirms h if and only if ***and to the extent that*** $P(h/e) > P(e)$
- ***N.B. 'Gives evidence in favour of' ≠ 'makes it more likely than not to be true'***