BIO3011
RESEARCH METHODS

Scientific Method: an introduction

Dr Alistair Hamilton

"If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts he shall end in certainties."

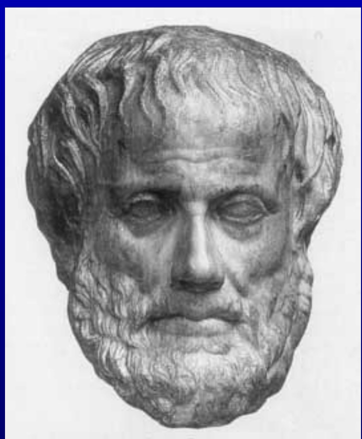
(Francis Bacon, 1605.)

Science is...

the systematic study of the nature and behaviour of the material and physical universe, based on observation, experiment and measurement

(Collins Concise Dictionary)

Aristotle (384-323 BC)



Aristotle (384-323 BC)

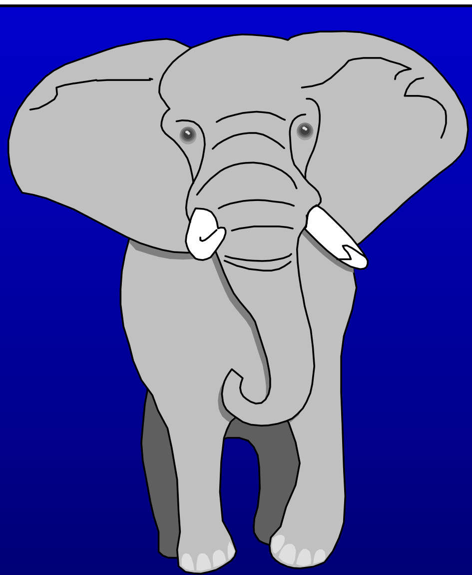
- rationalism - importance of reasoning from first principles
- teleology
- knowledge based on authority

Scientific revolution

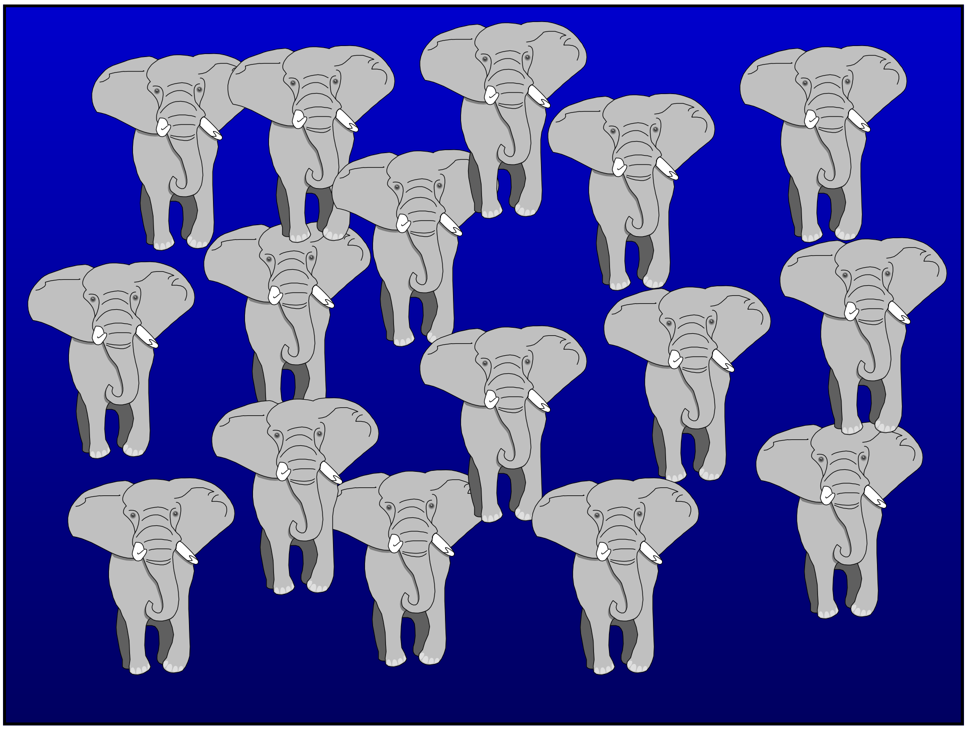
- late 16th early 17th century
- Francis Bacon (1561-1626)
- inductive reasoning

Inductive reasoning

process of reasoning from particular experiences or observations to general principles



An elephant at time t is observed to be grey



An elephant at time t was observed to be grey

I have observed many elephants under a wide variety of conditions, and they have all been grey

Therefore all elephants are grey

Singular statements

- that bird sounds an alarm call when disturbed
- that plant grows better when exposed to sunlight
- the litmus paper turns red when immersed in acetic acid

Universal statements

- birds in general sound an alarm call when a predator is seen
- plants require sunlight to grow
- acid turns litmus paper red

Principle of induction

'If a large number of A's have been observed under a wide variety of conditions, and if all those A's without exception posess property B, then all A's have property B.'

Chalmers (1999). This Thing Called Science.

Conditions of induction

- The number of observations must be very large
- The observations must be repeated under a variety of conditions
- No accepted observation should conflict with the derived universal law

Further developments

Further developments

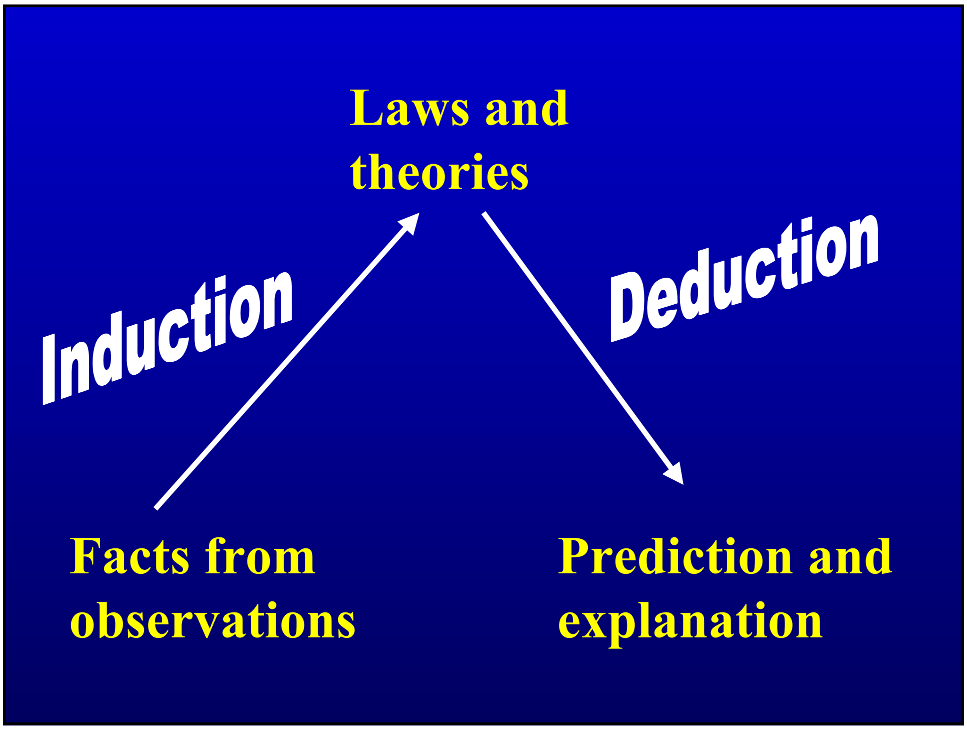
- Galileo - prediction and experimentation
- David Hume *et al* - empirical
- induction and deduction

Inductive reasoning

process of reasoning from particular experiences or observations to general principles

Deductive reasoning

to arrive at the necessary consequences, starting from admitted/established general principles



General form

LAWS and THEORIES
(via induction)

INITIAL CONDITIONS

PREDICTION and EXPLANATION
(via deduction)

Logical argument

All stats courses are boring	premise
This is a stats course	premise
This course is/will be boring	conclusion/ prediction

Logical argument

If the premises are true, and the argument is valid, then the conclusion MUST be true

Invalid argument

Many stats courses are boring

This is a stats course

This course is boring

False premise

All engineering courses are boring

This is an engineering course

This course is boring

Conditions of induction

- The number of observations must be very large
- The observations must be repeated under a variety of conditions
- No accepted observation should conflict with the derived universal law

Problems with induction

Large number of observations….

How many??

Wide variety of conditions….

What's important?

Observations…..

What about the 'unobservable'?

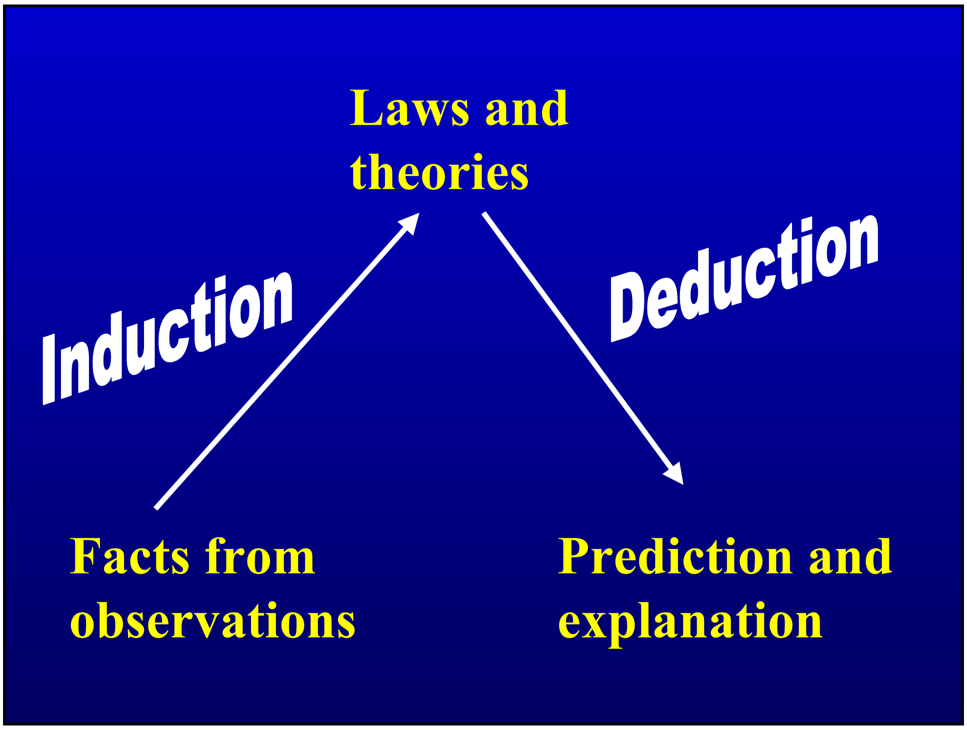
Problems with induction

the principle of induction worked on occasion 1

the principle of induction worked on occasion 2, etc etc

---

the principle of induction always works

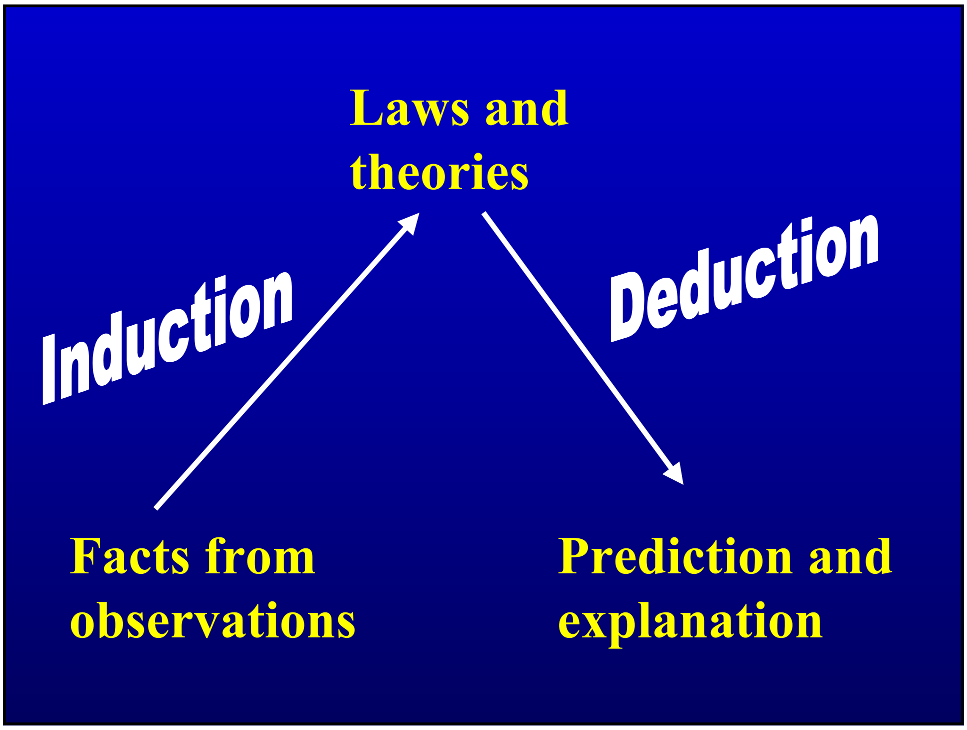


The scientific method and the role of statistics

Background reading

- **Chalmers, A.F. (1999).** *What is This Thing Called Science?* University of Queensland Press.

- **Ford, E.D. (2000).** *Scientific Method for Ecological Research.* CUP, Cambridge.



Karl Popper

Sir Karl Popper (1902-1994)

Karl Popper

- 1920's
- 'scientific' theories: empirical and via induction
- more facts = better?

'Current' theories

- Marx's theory of history
- Freud's psychoanalysis
- Adler's individual psychology
- Einstein's theory of relativity

The problems….

- very flexible - explanatory power with lots of verification
- 'could explain nothing because they could rule out nothing'
- note - not questioning the 'truth' of any theory

Popper's solution

- formulate knowledge into theories and then into hypotheses that can be rigorously tested
- these hypotheses MUST be falsifiable

Why the need to be falsifiable?

Falsifiable or not?

- elephants are either grey or another colour
- elephants are grey
- a bird may sound an alarm call when disturbed
- birds sound an alarm call when disturbed
- either it is raining or it isn't
- it never rains in Melbourne

More explanation the better….

- Mars moves in an ellipse around the sun
- Less information, restricted application,
- All planets move in an ellipse round the sun
- More information, wider application, easier to falsify, BETTER

Sulphur-crested cockatoos sound an alarm call when disturbed

All cockatoo species sound an alarm call when disturbed

Falsificationism

- A law/theory is conclusively falsifiable, but not conclusively verifiable
- all knowledge is permanently provisional (although for practical purposes……)
- 'the wrong view of science betrays itself in the craving to be right'

Popper

'our concern in the pursuit of knowledge is to get closer and closer to the truth, and we may even know we have made an advance, but we can never know if we have reached our goal'

Magee (1973). *Popper*.

In summary...

- we have a 'model' (theory), from which we can make predictions
- formulate logical hypothesis that can be falsified (tested)

What do we mean by 'hypothesis'?

- A set of propositions set out as an explanation
  - THEORY
- A conjecture, a new or unexplored idea
  - POSTULATE
- A statement that will be tested by investigation

Ford (2000)

The Null Hypothesis

- NH or H_0 usually formulated '….there is no difference….'
- Science is 'conservative'
- default option - accept H_0
- 'innocent till proved guilty…..'

Types of hypothesis

THEORY BUILDING

- test to detect *existence* of pattern
- test v randomness

THEORY TESTING

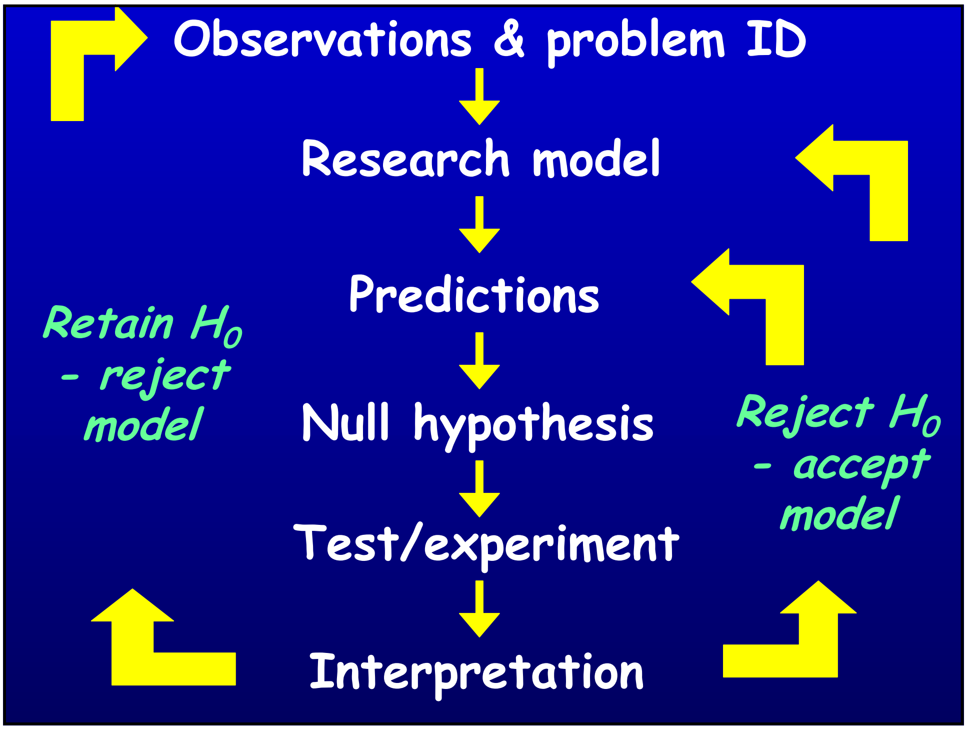
- test to detect *specific* pattern
- test v prediction

Building a research methodology
-
the hypothetico-deductive method

The hypothetico-deductive method

'The method of creating scientific theory from which the results already obtained could have been deduced and which also entails new predictions that can be verified or refuted by observation or experiment'

Ford (2000)



VARIATION IN TREE REGENERATION OBSERVED ACROSS THE RESERVE. LOTS OF DEER RUNNING AROUND.

THE DEER SEEM TO BE DISTRIBUTED UNEVENLY, AND THE VARYING GRAZING PRESSURE IS CAUSING THE OBSERVED VARIATION IN REGEN

EXCLUDING DEER FROM SOME AREAS WILL RESULT IN AN INCREASE IN REGEN IN THOSE AREAS

NH1.THERE IS NO DIFFERENCE IN REGEN BETWEEN FENCED AND UNFENCED AREAS

NH2. THERE IS MORE REGEN IN FENCED COMPARED TO UNFENCED AREAS

Why use hypothesis testing?

- appears to be objective and exact
- readily and easily used on PC's
- everyone else is doing it
- we are all taught to use it
- editors may demand it

Johnson (1999)

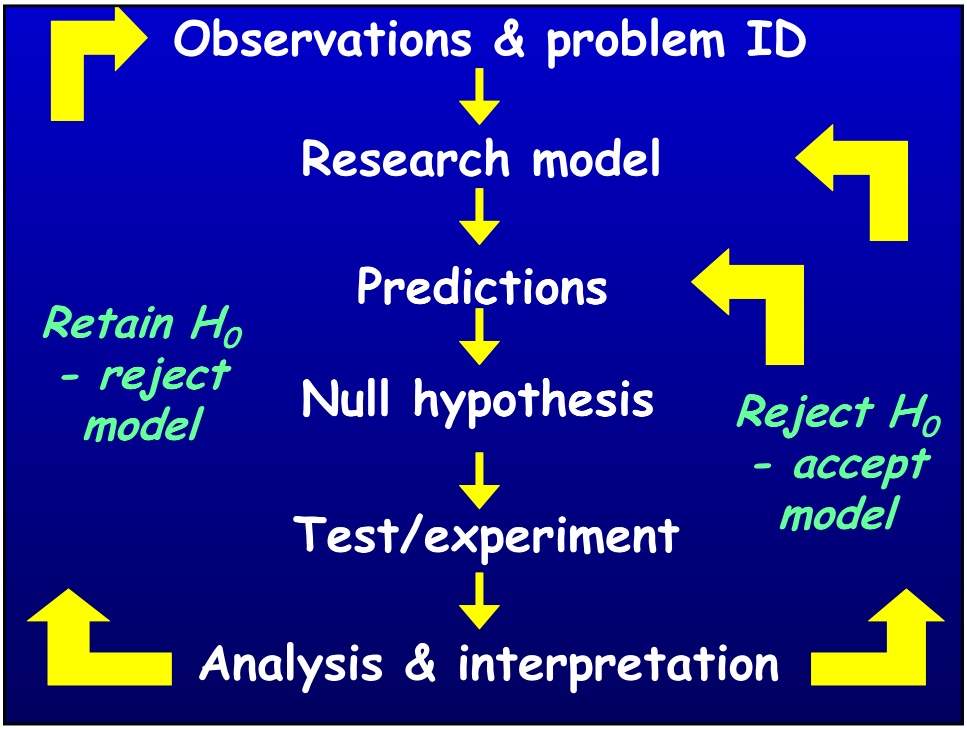
"the scientific approach has one characteristic that no other method of attaining knowledge has: self-correction. There are built-in checks all along the way to scientific knowledge."

Kerlinger (1973)

Statistical analysis can….

- identify real effects from random variation (separate signal from noise)
- describe distributions of data
- explore fit of data to models
- estimate and determine confidence intervals for parameters

Everything you always wanted to know about statistics, but were too afraid to ask…..



Definitions

- Population
  - biological: the individuals of a species that interact to maintain a homogenous gene pool
  - statistical: all individuals which are the subject of the study
- Unit or sample unit
  - individual items that make up a population

Definitions

- Variable
  - the characteristic (of the individual units) that we are interested in
- Observation
  - the recorded value of the variable of interest for one unit/individual (datum)
- Sample
  - a collection of observations from part of a population

Example 1
Number of plants in a quadrat

Population	all possible quadrats (in study area)
Unit	one quadrat
Sample	all quadrats actually surveyed
Observation	no of plants

Example 2
Mass of male possums

Population	all male possums (in study area)
Unit	one male possum
Sample	all male possums actually weighed
Observation	mass (kg)

Randomness and independence

- Sample units should be *randomly* drawn from a population (no bias), and should be *independent*

Types of data & scales

- discrete
- continuous

- nominal (categorical)
- ordinal
- interval
- ratio

Parameters & statistics

- Populations and parameters
- Samples and statistics

Pop mean μ **Sample mean $\bar{y}$**

Pop st dev σ **Sample st dev s**

Assumptions

- **Simplifying**
  - they are considered robust, what we 'know'
- **Explanatory**
  - what you are interested in / testing
- **Statistical**
  - these vary with the analysis being performed

Statistical analysis can….

- identify real effects from random variation (separate signal from noise)
- describe distributions of data
- explore fit of data to models
- estimate and determine confidence intervals for parameters

What a significance test *cannot* do is distinguish between

1. the null hypothesis is true
2. the alternative hypothesis is true

Too much emphasis on hypothesis testing?

What the P-value *doesn't* mean

- probability that results obtained were due to chance
- probability of getting the same result if experiment repeated
- probability that the null hypothesis is true

What P *does* mean

P is the probability of getting the observed or more extreme data, given that the null hypothesis is true

Significance levels

- 0.05 • significant (unlikely)
- 0.01 • highly significant (very unlikely)
- 0.001 • very highly significant (extremely unlikely)

Why use these values??

Non-significant results

- A non-significant result is a valuable result
- Null hypothesis cannot be 'proven' to be true
- NOT: '….results support/prove null hypothesis….'
- RATHER: '….results do not (yet) disprove/reject null hypothesis…'

Things to watch for….

- Confusing biological and statistical significance
- practicing the 'science of the bl**dy obvious!' (gratuitous hypothesis testing)

Reporting results

- reader must be able to critically assess your use of statistics and subsequent interpretation
- example from regression analysis: $R^2 = 0.86$, $F_{3,8} = 17$, $P < 0.001$

Possible outcomes from a test

Outcome of analysis	H_0 is (unknown to us) TRUE	FALSE
REJECT H_0	Type 1 error	Correct result
ACCEPT H_0	Correct result	Type 2 error

Errors in result...

- Type I (α) - rejecting a H_0 when it is actually true
  false positive
  FALSE 'ADVANCE'

- Type II (β) - accepting a H_0 when it is actually false
  false negative
  IGNORANCE

think about it…..

- If H_0 is true ('no difference'), and we reject it in error, we have made a Type I error
- probability of Type I set by us - α
- remember definition of P: 'the probability of getting the observed or more extreme data, given that the H_0 is true'
- α and β are inversely related

The power of a test

- is the probability of correctly rejecting the H_0 in favour of the alternative hypothesis H_A

Possible outcomes from a test

Outcome of analysis	H_0 is (unknown to us) TRUE	FALSE
REJECT H_0	Type 1 error	Correct result
ACCEPT H_0	Correct result	Type 2 error

The power of a test

- is the probability of correctly rejecting the H_0 in favour of the alternative hypothesis H_A
- power = 1 – β (probability of Type II error)
- is increased by increased sample size, decreased variance, increased effect size

Calculating power

- done *before* data collection
- to determine required sample size (need power and effect size)
- or to determine size of effect that can be detected (need power and sample size)

The 'ideal' statistical method should be

- CONSERVATIVE (low probability of making a Type I error)
- POWERFUL (low probability of making a Type II error)
- ROBUST (can deal with datasets that depart from assumptions)

Are there alternatives to hypothesis testing?

Confidence intervals

- 95 or 99 % CI
- One example of a procedure which, in 95% of cases, would include the true population mean

Precautionary principle 1

- Remember - conventional H_0 designed to err on side of caution
- H_0 - do we lack evidence that X affects Y?
- situations to err the other way?

Precautionary principle 2

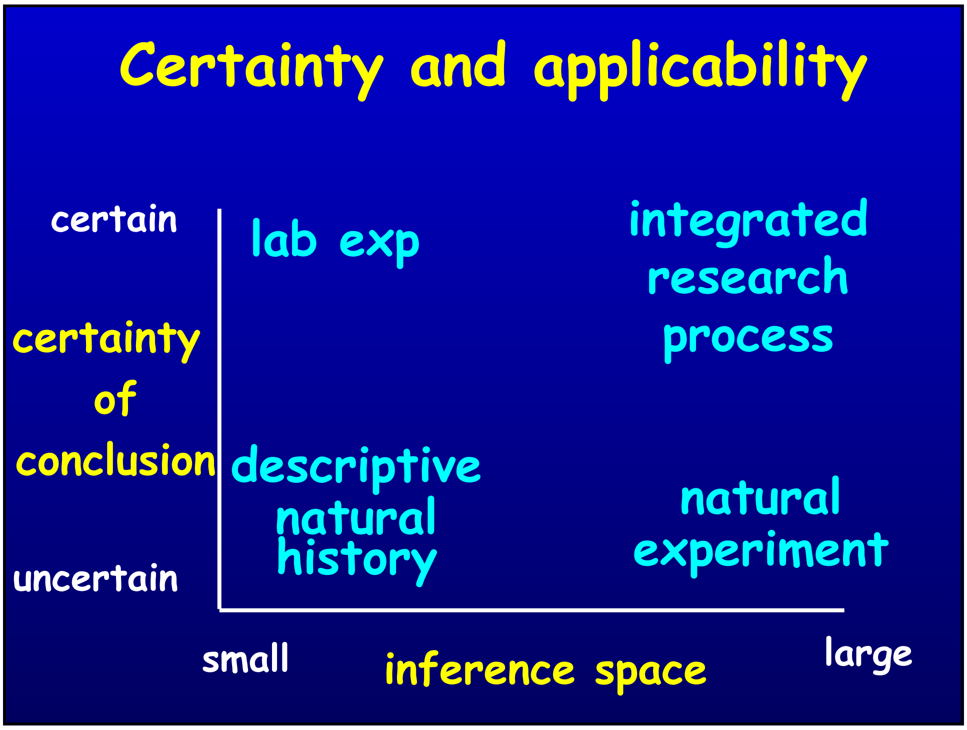
- if accepting H_0 not good enough, if we want positive evidence, ask 'do we have evidence that X *doesn't* affect Y?' (i.e. reverse burden of proof)
- consider equivalence and reverse tests

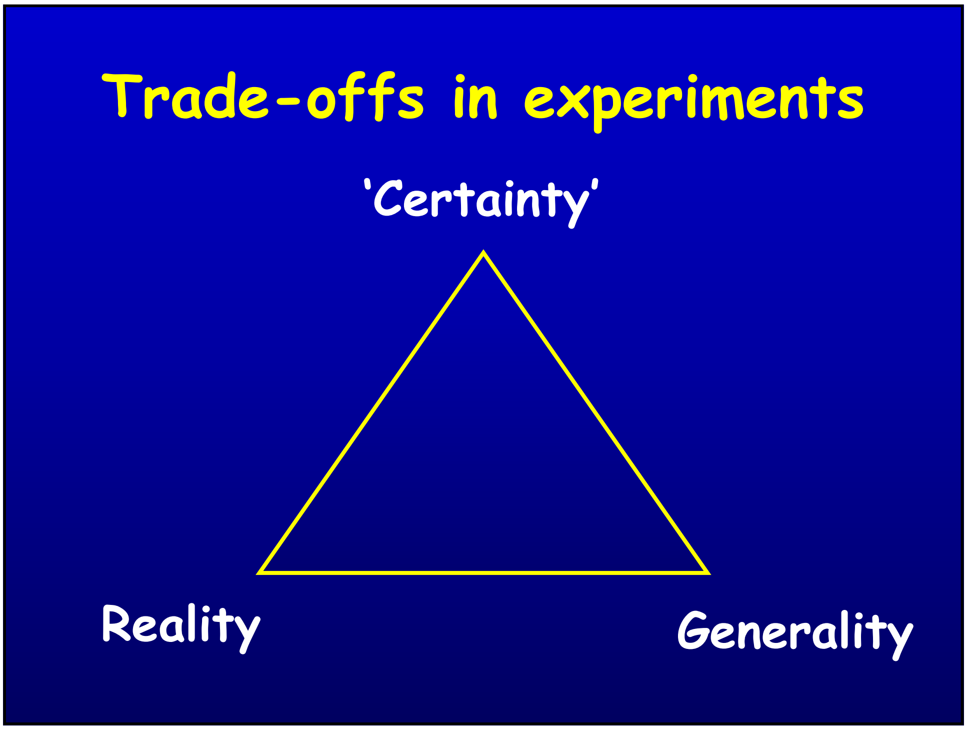
See Parkhurst (2001)

What scientists *actually* do

Types of experiment

- manipulative
- natural
- observational


Certainty and applicability
certain
certainty of conclusion
uncertain
lab exp
integrated research process
descriptive natural history
natural experiment
small
inference space
large


Trade-offs in experiments
'Certainty'
Reality
Generality

"science is a way of thinking more than it is a body of knowledge"

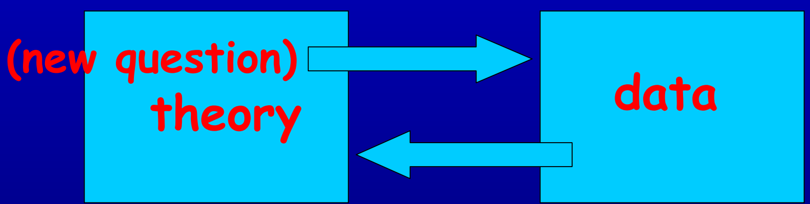
Carl Sagan (1979). *Broca's Brain*.

good science is asking the right questions in the right way

critical analysis

definition of

assessment of



developing new explanations and devising tests

seeing what nobody has seen before

creativity

Parkhurst, D.F. (2001). Statistical significance tests: equivalence and reverse tests should reduce misinterpretation. *BioScience* 51, 1051-1057

Johnson, D.H. (1999). The insignificance of statistical significance testing. *J. Wildlife Man.* 63, 763-772.

Hoenig, J.M & Heisey, D.M. (2001). The abuse of power: the pervasive fallacy of power calculations in data analysis. *The American Statistician* 55, 19-24.

Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54, 187-211.