UNIVERSITY OF HOHENHEIM

Scientific reasoning: Induction and deduction

3502-440 Methods of Scientific Working for Crop Science

WS 2024/2025

Prof. Dr. Karl Schmid
Institute of Plant Breeding, Seed Science and Population Genetics
University of Hohenheim

Motivation for reasoning

Arguments versus evidence

Need to differentiate arguments from evidence.

A first example of deduction

All plants contain DNA	Premise 1
Maize is a plant	Premise 2
Maize contains DNA	Conclusion

Another example

We randomly sample 400 out of 800 wheat plants	Premise 1
We found DNA in all 400 sampled wheat plants	Premise 2
The other 400 wheat plants have DNA, too	Conclusion

Is this a deductive argument?

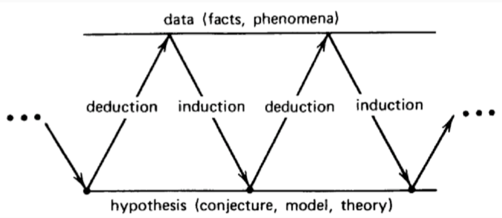
No, since the premise doesn't include the conclusion.

This is an inductive argument.

Key point:
how far can we extent: only 400 other wheat plants, all what plants all plants, all living systems?

Does deduction help us at all?

Iterative use of induction and deduction



Gauch, 2002

Model testing by deduction

1. Robert Grosseteste (Oxford, c. 1168-1280):
   Method of Verification and Falsification
   (a refinement of Aristotle's inductive-deductive method)
2. Given a model or hypothesis, deduct conclusions from it. They can then be tested in a controlled experiment
3. If the outcome from the experiment does not agree with the conclusion from the theory, it has to be discarded or refined
4. But: You can not verify a theory by empirical evidence, you can only falsify it (Karl Popper, 1902-1994)
5. Falsification of theories is the second main use of deduction
6. What good does positive evidence for a theory do then? ⇒ Induction

Example of a falsification

Hypothesis: All maize cobs are yellow

It is enough to find a single maize cob that is not yellow to falsify the hypothesis

Photo: Karl Schmid

The hypothesis has been falsified.

Using deduction: Mathematics

- Mathematics is almost entirely based on deduction
- Only a small set of assumptions called axioms are set to be true
- Every other conclusion is deducted from the axioms
- Aim for an axiom set as small as possible
- For nearly [1] all possible conclusions, you can theoretically deduct from the axioms whether they are true or not

[1] Gödel's first incompleteness theorem states that there are conclusions which cannot be shown to be either true nor false (or that there are contradictions)

Propositional logic

- In propositional logic, composite statements can be build up from propositions and operators
- Propositions: Simple statements like ”Plants have DNA”, ”Trees are plants” or ”Trees have DNA” (here symbolized by A, B and C). They can be true or false (nothing else)
- Operators (the most important ones):

Operator	and	or	implies	equals	not (negation)
Symbol	$\wedge$	$\vee$	$\rightarrow$	$=$	$\neg$

- $A \rightarrow B$ is often expressed as ”If A, then B”

Exercises

Translate these statements into formal logic:

- If plants have DNA and trees are plants, then trees have DNA:

$$(A \wedge B) \rightarrow C$$

- If trees have no DNA, then trees are no plants:

$$\neg C \rightarrow \neg B$$

- If trees have no DNA, then trees are no plants or plants have no DNA:

$$\neg C \rightarrow (\neg A \vee \neg B)$$

- If trees have DNA and trees are plants, then plants have DNA:

$$(C \wedge B) \rightarrow A$$

Note that this sentences are just composite statements and do not have a conclusion yet!

Logical arguments in propositional logic

- A logical argument is a list of statements (composite or simple), where the last statement is the conclusion, the rest are the premises. We will mark the conclusion by $\therefore$.
- An argument is valid if whenever all premises are true, also the conclusion is true
- For small arguments, this can be done by a truth table: To all propositions present in the argument, assign all combinations of true and false. Then, for each combination check whether the statements in the argument are true or false
- The argument is valid if every such combination that makes all premises true also makes the conclusion true

A truth table

```
julia> using TruthTables

julia> @truthtable (A ∧ B) --> C full=true
TruthTable
```

A	B	C	A ∧ B	A ∧ B --> C
true	true	true	true	true
true	true	false	true	false
true	false	true	false	true
true	false	false	false	true
false	true	true	false	true
false	true	false	false	true
false	false	true	false	true
false	false	false	false	true

How to work with propositional logic

Check whether you are making a deductive or non-deductive argument:

- Deductive argument: If premises are true the conclusions are inevitably true
- Nondeductive argument: Conclusion is beyond premises → Induction

In real life: Premises are often not clear or well definable to construct a deductive argument

How to work with propositional logic

For a deductive argument, split a complex argument into simpler and valid rules or equalities

Examples (there are many more) for inference rules/equalities (p, q are statements):

Modus ponens	$p, p \to q \therefore q$
Modus tollens	$\neg q, p \to q \therefore \neg p$
Modus tollens (as equality)	$(p \to q) = (\neg q \to \neg p)$
De Morgan's Rule 1	$\neg(p \wedge q) = \neg q \vee \neg p$
De Morgan's Rule 2	$\neg(p \vee q) = \neg q \wedge \neg p$

Modus ponens

Modus ponens: The mode of putting: put p, get q

- If *p* then *q*.
- *p*.
- Therefore, *q*.

Modus tollens

Modus tollens: The mode of taking: take q, take p

- If *p* then *q*.
- Not *q*.
- Therefore, not *p*.

Example of *modus tollens*

Sherlock Holmes:

> *A dog was kept in the stables, and yet, though someone had been in and had fetched out a horse, [the dog] had not barked.... Obviously the ... visitor was someone whom the dog knew well.*

Put as *modus tollens*:

- If the visitor were a stranger (*s*), then the dog would have barked (*b*)
- The dog did not bark.
- Therefore, the visitor was not a stranger.

- if s then *b*.
- Not-*b*.
- Therefore, not-s.

Motivation: What is scientific reasoning?

What is deduction?

Structure of deductive arguments

Deduction fallacies, blunders and failures

What is induction?

Logical fallacies

What are logical fallacies?

- Bad logic
- Incorrect arguments
- Weapons in a philosophical (or scientific) argument
- 232 Types of logical fallacies were described

Fallacies

- **Definition:** Incorrect reasoning in argumentation. The result is a misconception.
- Occurs in rhethorics, and in logics.
- Different types of fallacies: Logical, material and verbal fallacies.
- Fallacies were already described by Aristotle
- Medieval School of Scholastics identified numerous types of fallacies
- Use and abuse of fallacies in rhethorics, science, business and politics

Blunders in deduction

Often deductions are misused or wrong deductions are made (accidently and on purpose). Here are the most common blunders:

Invalid reversion of modus ponens

Wholefood shop keeper: ”We know that selling GMO ingredients will drive away customers. I have lost customers, so my suppliers must have mixed GMO ingredients in the products”

$B, A \to B \therefore A$ is not valid!

Invalid ”negation” of modus ponens

”If I get a flu shot, I won’t get the flu. I didn’t get a flu shot, so I will get the flu”

$\neg A, A \to B \therefore \neg B$ is not valid

Common fallacies

Composition

- A property of parts is applied to the whole
- “Sodium and chlorine are poisonous; table salt is sodium chlorine; Therefore, table salt is poisonous.

Division

- Apply a property of a whole to individual parts.
- “Dogs are common; albino spaniels are dogs; therefore albino spaniels are common.”

Common fallacies

Circular logic

- Assumes what it intends to prove: $p \therefore p$
- ”Assumption: All plants have DNA. Douglas firs have DNA since we know that douglas firs are plants and plants have DNA.”
- May seem plausible, since often the assumption of p is obscured

Another blunder

- ”We have to either cut money for police or for universities. Since we can't cut the money for police, we have to cut universities' money”
- Is this a blunder? It's a valid logic argument.
- Though $A \vee B, \neg A \therefore B$ is valid, the first premise may be wrongly chosen (it could be $A \vee B \vee C \vee \ldots$). But changing this renders the argument invalid!
- So maybe we could cut money somewhere else. Then cutting the universities' budget wouldn't be a valid conclusion any more.
- This blunder is called false dilemma and is one of the more common blunders. It is often used intentional.
- The ”opposite” blunder: Believe there are more options available than there actually are

Intentional logical fallaciess

Straw man argument

- Present the opponents position in a simplistic manner (the straw man), then contradict the simplistic/distorted arguments
- **Real position of A**: ”We have to be careful with new GMO plants and should test thoroughly for harmful effects before cultivating them”.
- **Constructed straw man position**: ”We should not cultivate any GMO plants since they may have very minor side effects”
- **Attack the straw man**: ”A is killing people since A wants to forbid planting GMO plants because of minor side effects”
- What would be the correct way: present the strongest case of the opponent and check whether you can contradict the strongest case
- **Very frequently used tactic!**

Intentional logical fallacies

Argument from ignorance

- Drive opponents to accept my argument unless they can find a better argument to the contrary
- “We cannot prove that this pesticide is safe, so we must assume it is dangerous and outlaw its use”
- Failure to reject a hypothesis is often taken as a proof of the hypothesis
- Implicit argument: “Give me a better argument or either accept my argument”
- Especially powerful if the discussed question cannot be answered completely (climate change, darwinism vs. creationism)

Other blunders in research

- Accidental blunders
- Unexamined presupposition, bad data and invalid logic → wrong conclusions
- Sometimes lack of will (sloppiness) or lack of competence

These fallacies are very frequent in science!

Example: We want to show that a certain variety has better yield in a certain environment than other varieties.

Our hypothesis is that it is resistant against a common pathogen in this environment.

We show that the variety actually has this resistance, thus claiming we have shown that the hypothesis is correct.

Using deduction: Distinguishing crop varieties

Look at the variety description of the winter barley variety ”Tout en val” from the Community Plant Variety Office (CPVO) of the European Union.

- Where is the principle of deduction applied here?
- How can you distinguish ”Tout en val” from the similar variety ”Jolival”?
- For a single plant, we want to check whether it may belong to the variety ”Tout en val”. Is it easier to find out that it belongs or that it does not belong?
- Is distinguishing between varieties completely deductive?

Summary

- Deduction reasons from general to specific
- Conclusions from valid deductive arguments are 100% true
- Uses of deduction: Deduct hypotheses from theoretical models or falsify theories
- The validity of logical arguments can be checked with a truth table
- Deduction is often applied incorrectly. This misuse results in unjustified conclusions.

Further reading

- Weston (2009) *A rulebook for arguments.* 4th edition. Hackett Publishing Company. 88p. Excellent short and very accessible introduction into making arguments. See Chapter VI.
- Gauch (2002) *Scientific Method in Practice.* Cambridge University Press. Chapter 5.
- Priest (2000) *Logic - A Very Short Introduction* Oxford University Press.

Review questions

- Why is deduction important for doing science?
- Can you verify a deductive argument by an experiment?
- We have the hypothesis that a certain disease is caused by either a deleterious recessive allele at locus A or at locus B. The disease is very rare. We're not really sure whether the hypothesis is true. Design an experiment to try to falsify this hypothesis.
- Write a truth table for $(A \wedge B) \rightarrow C \therefore (C \wedge B) \rightarrow A$ to show whether it's a valid logical argument
- Find new examples for the common deductive blunders described in the lecture

A first example of induction

- Assume a population fo 800 wheat plants
- 400 plants were tested whether they contain DNA

Analysis by induction:

We found DNA in 400 of 800 observed wheat plants	Premise 1
The other 400 wheat plants have DNA, too	Conclusion

Properties of induction

- An inductive argument consists of premises and conclusions
- The premises do not contain the conclusions
- If the premises are true, the conclusions are not necessarily true. However, in a reasonable inductive argument the conclusions will be true with high probability
- Induction reasons from specific to general. In other words, it reasons from the observed to the unobserved

Some modes of induction

(a) Induction from sample to population

(b) Induction from history to presence/future

(c) From entities to similar entities

For a: Statistics is required to test how much the sample resembles the population For b: Prediction is required: Which model is the best to predict the future from the past For c: Which entities are comparable?

The inheritance of a gene in maize should be in the same manner than in rice

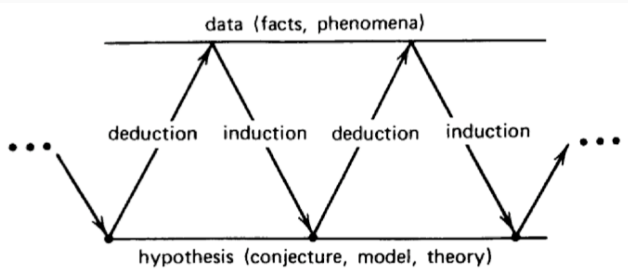
The importance of induction

- It is easier, cheaper or only possible to analyse a sample instead of a whole population of objects
- Induction is the basis of any prediction for future events
- It lays the foundation for models, hypotheses or theories. From these, deductive inference for unobserved objects is possible
  ⇒ Iteration of induction and deduction)

The difference between induction and deduction

	Deduction	Induction
Conclusion	Contained in premises	Extension of premises
Truth	Certain	Probable
Direction	General Model → Reality	Reality → General Model

Aristotle's key contribution: Inductive-deductive methods



Induction versus deduction

The iteration of induction and deduction can be nicely demonstrated with the Hardy-Weinberg-Equilibrium.

Problem 1. Assume that a very large plant population is in Hardy-Weinberg equilibrium at a locus. We sample 100 individuals.
$\rightarrow$ What genotype frequencies at this locus can we expect in the sample?
$\Rightarrow$ Deduction with the Hardy-Weinberg-Formula

Problem 2. We observe the genotype frequencies in a sample of 100 plants.
$\rightarrow$ What are the genotype frequencies in the population?
$\Rightarrow$ Induction: Is the population in Hardy-Weinberg equilibrium?

These types of quantitative inference (both deductive and inductive) need a mathematical model describing the relationship between sample and population

Presuppositions of induction

- **Uniformity of nature (UN):** Universe is governed by general laws. Objects which are "identical" should act the same way under the same circumstances.
- **Parsimony or law of limited variety of nature:** Things represent one another in relevant aspects and are not unlimited in their variety. Induction is *within* classes of objects

Criticism of induction (Hume's problem, 18th century)
UN can neither be justified by induction nor deduction

$\Rightarrow$ Still, induction works good enough in general

Uniformity of nature refers both to the perception and the existence of nature. On a more mundane level, much of plant breeding and plant cultivation tests the rule for UN. On a very basic level it likely may be true, but not on this higher level.

Similar things behave similar: Atmosphere of Earth vs. Atmosphere on Mars Agriculture: Optimal plant cultivation in tropics vs. optimal cultivation in Europe

Induction and mathematics

- Quality of an inductive argument: If the premises are true, the conclusion should also be true with high probability
- What is probability? How can you measure its value?
- We use a mathematical model which describes both premises and conclusion (think about sample and population). From this model, we deduce the probability of the conclusion (or at least estimate or find boundaries)
- This branch of mathematics is called statistics
- Applied induction *is* statistics
- Problem: How to choose the mathematical model? Here, we already have to make an inductive step of choosing a model given our knowledge of the premises and conclusion.

Yes or no were in deductive logic binary traits.

The scope of statistics

Statistics has two major branches (we assume here induction from sample to population):

- Estimation of parameters
- Testing of models

Estimation in statistics

Definition: Estimate parameters of the population using parameter values from the sample.

Example: Estimate mean 1000 grain weight in a population of a wheat elite line by measuring the trait in a sample and computing an estimate from the data.

Quality criteria:

- **Confidence interval:** Give a range of parameter values so that the correct population parameter lies in this range with a certain (high) probability. Small intervals are a sign of a good estimation.
- **Unbiasedness:** If you repeat the estimation in a lot of samples, the mean value of all estimations is the population value
- **Consistence:** If you enlarge your sample, the estimation gets better (less estimation error)

Testing in statistics

Definition:

- Develop a model or hypothesis about a phenomenon in a population
- Get a sample for the population
- Use data from sample and a procedure to decide whether a hypothesis or model is correct

There are two main schools of thinking how to come up with such a procedure

- **Frequentist school:** Test only a single hypothesis and choose an inference rule that the decision based on the sample will be correct for many other samples drawn from the population
- **Bayesian school:** Among multiple hypotheses, choose the one with the highest probability given the observed data

Practical problems with induction

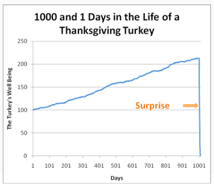
Imagine an urn with either 5000 green and 5000 red balls or 4800 green and 5200 red balls.

- **Small samples:** If samples are too small, we can not conclude with high probability. Drawing just two balls out of the urn does not help deciding which possibility is true.
- **Biased sampling:** Samples are drawn in a specific manner, e.g. draw only red balls. Such samples have to be analysed carefully and with respect to the drawing rule, since they are most likely not representative for the population (in contrast to a random sampling). This is often done on purpose to find false evidence for a hypothesis (Cherry-picking).

Practical problems with induction

Rare events:

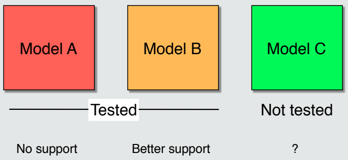
- Very rare events may not occur even in very big samples
- The probability of such events can not be calculated from a sample
- This is a problem if the rare events violate the conclusions
- Especially: We can not predict the future from the past



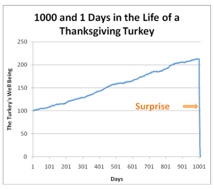
Practical problems with induction

Type M (Model) error:

- Use a sample of a population to distinguish between several false hypotheses.
- Data may support better the "less false" hypothesis which does not make it more true.



"Essentially, all models are wrong, but some are useful" (Statistician G.E.P. Box)



Summary

- Induction derives general claims from a sample of observations
- The general claim is not true *per se*, but only likely true (with high probability)
- Basically, the scientific method is built from the interplay of deduction and incduction
- Statistics is the application/quantification of induction

Further reading

- Weston (2009) *A rulebook for arguments.* 4th edition. Hackett Publishing Company. 88p. Excellent short and very accessible introduction into making arguments. See Chapters II, III, IV, V, VII.
- Gauch (2002) *Scientific Method in Practice.* Cambridge University Press. Chapter 7.

Review questions

- What are the key differences between inductive and deductive arguments?
- Why is the uniformity of nature assumption so important in the use of induction?
- Do you know some examples in agriculture in which the assumption of a uniformity of nature is violated?
- What is the use of statistics in inductive arguments?

References i

Gauch, H. G. (2003). *Scientific Method in Practice.* Cambridge University Press.