Probability

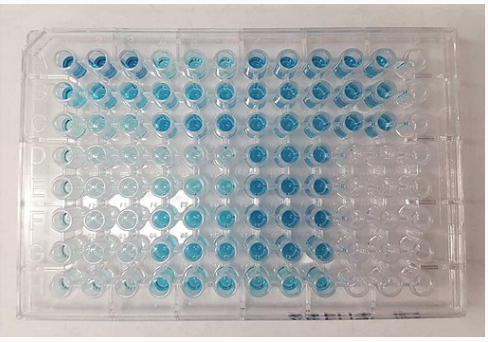
3502-440 Methods of Scientific Working for Crop Science

WS 2024/2025

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

Motivation

Enzyme-linked immunosorbent assay (ELISA)



Definition of Probability Theory

Formal definition of probability theory

Bayes' Theorem

Common errors in probability theory

Probability and Statistics

- Probability theory is derived from deductive logic
- Statistics is derived from inductive logic
- Both are tools for inference
- Many decisions are based on more or less conscious probabilistic arguments
- Misuse and abuse of probability and statistics is common...
- ...even in science

Conceptual ideas of probability

- Objective or physical probability: Chance of an event occurring
- Subjective, personal or epistemic probability: Degree of belief in a proposition warranted by experience

Different meanings of probability

(1) A fair coin has a probability of 0.5 of heads, and likewise 0.5 of tails; so the probability of tossing two heads in a row is 0.25.
(2) There is a 10% probability of rain tomorrow.
(3) There is a 10% probability of rain tomorrow according to the weather forecast.
(4) Fortunately there is only a 5% probability that her tumor is malignant, but this will not be known for certain until the surgery is done next week.
(5) Smith has a greater probability of winning the election than does Jones.
(6) I believe that there is a 75% probability that she will want to go out for dinner tonight.
(7) I left my umbrella at home today because the forecast called for only a 1% probability of rain.
(8) Among 100 patients in a clinical trial given drug A, 83 recovered, whereas among 100 other patients given drug B, only 11 recovered; so new patients will have a higher probability of recovery if treated with drug A.

Gauch (2003)

Interpretation of statements

- (1): Deductive; apodictive; self-evidently true
- (2): Empirical statement without evidence
- (3): Empirical statement with source of evidence (but not evidence itself)
- (4): Probability refers to knowledge not the fact itself; probability may change with additional knowledge
- (5): Relative probability, but information is incomplete
- (6): Subjective estimation of probability, no evidence given
- (7): Probability as basis for decisions and actions; value is part of cost/benefit analysis
- (8): Induction: Comparison of two singular observations, derive general conclusion

Aspects in the definition of probability

- Events versus beliefs
- Repeatable versus single events
- Expression in exact numbers versus inexact beliefs
- A function of one argument (event) or two arguments (event and evidence)
- Combine or not combine old and new information?
- Effects of ignorance versus knowledge on probability statements
- Combination of theoretical and empirical aspects of probability
- Connection between deductive and inductive applications?
- Single theory of probability for all situations possible?

Fundamental requirements of probability

- **Generality:** Works for all cases and persons
- **Impartiality:** Must be fair to all hypotheses

Eight general rules for a concept of probability

- Explicit
- Coherent / self consistent
- Practical: Experiments should be possible
- Revisable
- Empirical: Conclusions must be dominated by evidence
- Parsimonious: Number of axioms should be small
- Human: Compatible with humaness and imperfection
- Not perfectionistic: Take care of experimental error

First candidate: Relative frequencies

Consider a (finite) set of objects with different traits $A, B, \ldots$.
Define the relative frequency of trait A as

$$\mathrm{fr}(A) = \frac{\text{number of objects in the set with trait } A}{\text{total number of objects in the set}}$$

The relative frequency of objects with a specific trait observed so far should be a predictor for the number of objects with the trait among a set of newly observed objects.

Assumption: UN principle (uniformity of nature)

Properties of relative frequencies

- Relative frequencies are values ≥ 0
- Relative frequency of objects with any trait is 1
- $\mathrm{fr}(A \text{ or } B) = \mathrm{fr}(A) + \mathrm{fr}(B)$ for mutually exclusive traits A and B

This is the basis of the concept of probability

Axiomatic definition of probability

Probability axioms from Kolmogoroff (1930's):

- (1) Probability of getting a particular outcome is expressed as value ≥ 0
- (2) Probability of getting any possible outcome is 1
- (3) Probability of getting either of two mutually exclusive outcomes equals the sum of the probabilities of these outcomes.

Axiomatic definition of probability - Example

Toss two coins. We set the probability for each of the four outcomes $\frac{1}{4}$.

- The probability of getting two heads is $1/4$
- The probability of getting any outcome, namely two heads or two tails or else one heads and one tails, is 1.
- The probability of getting two heads, or else one heads and one tails is $1/4 + 1/2 = 3/4$

Do these statements make sense?

- (1) The probability of an ideal fair coin landing heads is $1/2$
- (2) The probability of an actual fair coin landing heads is nearly $1/2$
- (3) The probability that my belief "The fair coin will land heads" will be true is $1/2$.

Apply concepts to

- abstract entities
- actual events
- personal beliefs

$\rightarrow$ Theory serves all ordinary and scientific applications

Definition of Probability Theory

Formal definition of probability theory

Bayes' Theorem

Common errors in probability theory

Formal definition of probability axioms

First, some formal definitions:

- Two sets $A = \{1, 2\}$ and $B = \{2, 3, 4\}$
- Union: $A \cup B = \{1, 2, 3, 4\}$
- Intersection: $A \cap B = \{2\}$
- Subset: $\{2\}$ of $\{1, 2\}$ but not of $\{3, 4\}$
- Mutually exclusive: $C = \{2\}$ and $D = \{3, 4\}$. Formally, mutual exclusivity is written as $C \cap D = \emptyset$ (empty set)
- Complement: All possible outcomes form the set Ω. Let A be a subset of Ω. The complement $\neg A$ of A consists of all outcomes in Ω which are not in A.

Kolmogoroff's axioms in formal language

Let Ω be a (finite) set of outcomes and A, B subsets. P is a probability measure if

- $0 \leq P(A)$
- $P(\Omega) = 1$
- If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

Advantages of (formal) probability theory:

- Thoughts have common structure and predictability
- Thinking is public $\rightarrow$ common understanding is possible

Conditional probability

- Probability of H given D: $\mathrm{P}(H|D)$
- H: Hypothesis; D: Data (both are subsets of outcomes, $P(D) > 0$)
- Definition of conditional probability:

$$P(H|D) = \frac{P(H \cap D)}{P(D)}$$

- Interpretation: Knowing the actual outcome of an experiment is in D, what is the probability that an outcome of H has happened?

Example of conditional probability

Assume that we search for a genebank accession resistant against a pathogen.

For this, we screen 100 accessions for disease symptoms after infecting them with the pathogen

Goal: Calculate conditional probability of no symptoms given an accession is resistant

	No symptoms	Symptoms	Marginal
Resistant	4	1	5
Susceptible	5	90	95
Marginal	9	91	100

Example of conditional probability

	No symptoms	Symptoms	Marginal
Resistant	4	1	5
Susceptible	5	90	95
Marginal	9	91	100

Unconditional probability that accession without symptom is resistant: $\frac{9}{100} = 0.09$

Conditional probability of accession:

- $P(\text{No Symptom} \cap \text{Resistant}) = \frac{4}{100} = 0.04$
- $P(\text{Resistant}) = \frac{5}{100} = 0.05$

Then:

$$P(\text{No Symptom}|\text{Resistant}) = \frac{P(\text{No Symptom} \cap \text{Resistant})}{P(\text{Resistant})} = \frac{0.04}{0.05} = 0.8$$

Probability in action: Deduction of conclusions

Conclusion 1: $P(\neg A) = 1 - P(A)$

Proof: A and $\neg A$ are mutually exclusive ($A \cap \neg A = \emptyset$) and $A \cup \neg A = \Omega$.

$$1 = P(\Omega) = P(A) + P(\neg A)$$

Probability in action: Deduction of conclusions

Conclusion 2: $P(B) = P(B|A)P(A) + P(B|\neg A)P(\neg A)$

Proof: $B = \underbrace{(A \cap B)}_{:=C} \cup \underbrace{(\neg A \cap B)}_{:=D}$. C and D are mutually exclusive, so

$$P(B) = P(C) + P(D) = P(B|A)P(A) + P(B|\neg A)P(\neg A),$$

since $P(C) = P(A \cap B) = P(B \cap A) = P(B|A)P(A)$ by the definition of conditional probability and analogously $P(D) = P(B|\neg A)P(\neg A)$

All theorems and techniques in probability theory are deducted from the probability axioms in this way

Bayes' Theorem

- Reverend Thomas Bayes (1702 - 1761)
- Describes the degree of confidence given prior evidence

Simple form of Bayes' Theorem:

$P(H|D) \propto P(D|H) \times P(H)$ or

Posterior $\propto$ Likelihood $\times$ Prior

- Bayes' Theorem, complete form

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} = \frac{P(D|H)P(H)}{P(D|H)P(H) + P(D|\neg H)P(\neg H)}$$

A proof of Bayes' Theorem i

- Theorem: $P(H|D) = \frac{P(D|H) \times P(H)}{P(D)}$
- Rearrange definition of conditional probability:

$$P(H \cap D) = P(H|D) \times P(D)$$

- Second application with H and D reversed:

$$P(D \cap H) = P(D|H) \times P(H)$$

- $H \cap D$ and $D \cap H$ are same set, so

$$P(H \cap D) = P(D \cap H)$$

- Equate the formulas to obtain

$$P(H|D) \times P(D) = P(D|H) \times P(H)$$

A proof of Bayes' Theorem ii

- Ignore normalizing constant, $P(D)$ and substitute proportionality for equality, we get Bayes' theorem in simple form
- Use Conclusion 2 to get the complete form

Important terms in Bayes' Theorem

$$P(H|D) \propto P(D|H) \times P(H)$$

- $P(H)$: Prior probability or just prior
- $P(D|H)$: Likelihood. Data's impact on the probabilities of hypotheses
- $P(H|D)$: Posterior probability or posterior

  Bayes' theorem solves the inverse or inductive problem of calculating the probability of a hypothesis given some data $P(H|D)$ from the probability of some data given a hypothesis $(D|H)$

Bayes' Theorem with competing, mutually exclusive hypotheses

- Formula for two hypotheses ($H_2 = \neg H_1$):

$$P(H_1|D) = \frac{P(D|H_1) \times P(H_1)}{(P(D|H_1) \times P(H_1)) + (P(D|H_2) \times P(H_2))}$$

- Formula for multiple hypotheses:

$$P(H_1|D) = \frac{P(D|H_1) \times P(H_1)}{\sum_{i=1}^{n}(P(D|H_i) \times P(H_i))}$$

- The ratio form compares hypotheses in forms of ratios or odds:

$$\frac{P(H_1|D)}{P(H_2|D)} = \frac{P(D|H_1)}{P(D|H_2)} \times \frac{P(H_1)}{P(H_2)}$$

Exercising...

Consider a biallelic locus with alleles A_1, A_2 in a diploid population which consists of two subpopulations of equal size.

- In Subpopulation 1, 50 % of the (sub)population is heterozygous (of genotype A_1A_2) at the observed locus
- In Subpopulation 2, 25 % of the (sub)population is heterozygous

Pick an individual at random from the population. It's heterozygous at the observed locus. What's the probability that it originates from Subpopulation 1?

Solution i

Define

- $H :=$ picked individual is from Subpopulation 1
- $D :=$ picked individual is heterozygous

We already know

- $P(H) = P(\neg H) = 0.5$, since both subpopulations are equal in size
- $P(D|H)$ is the probability that, if we pick from Subpopulation 1, we pick a heterozygous individual, which is the frequency of heterozygotes in SP1, hence $P(D|H) = 0.5$. Analogously $P(D|\neg H) = 0.25$, since this is the the probability that, if we pick from Subpopulation 2, we pick a heterozygous individual

Solution ii

We want to compute $P(H|D)$. Using Bayes' Theorem,

$$P(H|D) = \frac{P(D|H)P(H)}{P(D|H)P(H) + P(D|\neg H)P(\neg H)} = \frac{0.5 \cdot 0.5}{0.5 \cdot 0.5 + 0.25 \cdot 0.5} = \frac{2}{3}$$

Common errors in probability theory

- Ignored prior
- Ignored condition

Ignored prior i

You have a test for some rare disease which occurs by chance in 1 in every 100,000 trees (e.g., mango *Mangifera indica*). The test is fairly reliable; if a tree has the disease it will correctly say so with probability 0.95; if a tree does not have the disease, the test will wrongly say it does with probability 0.005. If the test says your tree will suffer from the disease, what is the probability that this is a correct diagnosis? Is it 0.95?← blunder

Solution:

- Probability that your tree has the disease, $P(D) = 0.00001$
- Probability that your tree does not have the disease, $P(\neg D) = 0.99999$

Ignored prior ii

- Probability of positive test when the tree has the disease, $P(T|D) = 0.95$
- Probability of positive test when the tree is healthy, $P(T|\neg D) = 0.005$
- Probability of disease given positive test result:

$$
\begin{aligned}
P(D|T) &= \frac{P(T|D) \times P(D)}{P(T|D) \times P(D) + P(T|\neg D) \times P(\neg D)} \\
&= \frac{0.95 \times 0.00001}{(0.95 \times 0.00001) + (0.005 \times 0.99999)} \\
&\approx 0.002
\end{aligned}
$$

Keep in mind: $P(T|D) \neq P(D|T)$!

Ignored condition i

Consider:

- A plant may be resistant to two pathogens P_1 and P_2
- $P(P_1\text{resistance}) = P(P_2\text{resistance}) = 0.5$
- Four possibilities of resistances: none, $P_1 (= (P_1, \neg P_2))$, $P_2 (= (P_2, \neg P_1))$, (P_1, P_2)
- We assume that resistance to each pathogen is independent from the other resistance: $P(P_1, P_2) = 0.5^2 = 0.25 = P(P_1) = P(P_2) = P(\text{none})$
- Question 1: What is the probability a plant has both resistances? Answer: $Prob = 1/4$

Ignored condition ii

- Question 2: A plant has one of the resistances. What is the probability that it also has the other resistance? Is it again $\frac{1}{4}$? Is it $\frac{1}{2}$? $\leftarrow$ blunder
  Answer: no resistance is not possible; hence: $\underline{P_1, P_2}$, P_1, P_2
  $\rightarrow Prob = 1/3$

Formal derivation using conditional probability

- X: There is at least one resistance
- Y: $P_1 P_2$ resistance
- $P(X) = 3/4$ (3 of 4 combinations have a resistance)
- $P(X \cap Y) = 1/4$ (only 1 of 4 has both resistances)
- Hence: $P(Y|X) = P(X \cap Y)/P(X) = (1/4)/(3/4) = 1/3$

Summary

- Probability is useful for deduction and induction
- Formal definition of probability theory rests on few axioms
- Theorems can be derived from the axioms
- Bayes' Theorem measures confidence in hypothesis given data
- Blunders in probability theory: Incorrect use of prior; neglect of prior information

Further reading

- Any introductiory textbook on statistics and probability
- This chapter is based on Gauch (2003), Chapter 6. For a more concise introduction, see Hugh G. Gauch (2012), Chapter 8.

Review questions

1. Could Bayes' theorem be useful for inductive procedures? How/why?
2. A test for skabies[1] predicts 98% correctly to healthy people that they are free of skabies, and 98% correctly to sick people that they have skabies. We assume that 1 in 1000 inhabitants in our city have skabies. The test is positive for you.
   a) What is the probability that you have skabies given the positive test?
   b) What is the probability if you use a flat prior, i.e. a probability of skabies = 0.5?
3. Find examples for the blunders ”Ignored prior” and ”Ignored condition” surrounding Bayes' theorem

[1]Skabies is a contageous skin disease that is caused by mites; In German: Krätze

References i

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

Hugh G. Gauch, J. (2012). *Scientific Method in Brief*. Cambridge University Press.