Statistics in Geophysics: Introduction and Probability Theory

Steffen Unkel

Department of Statistics
Ludwig-Maximilians-University Munich, Germany

What is Statistics?

- "*Statistics is the discipline concerned with the study of variability, with the study of uncertainty, and with the study of decision-making in the face of uncertainty.*" (Lindsay, et al. (2004): A report on the future of Statistics, *Statistical Science*, Vol. 19, p. 388)

- Statistics is commonly divided into two broad areas:
  - Descriptive Statistics
  - Inferential Statistics

- The descriptive side of statistics pertains to the organization and summarization of data.

- Inferential statistics consists of methods used to draw conclusions regarding underlying processes that generate the data (population), by examining only a part of the whole (sample).

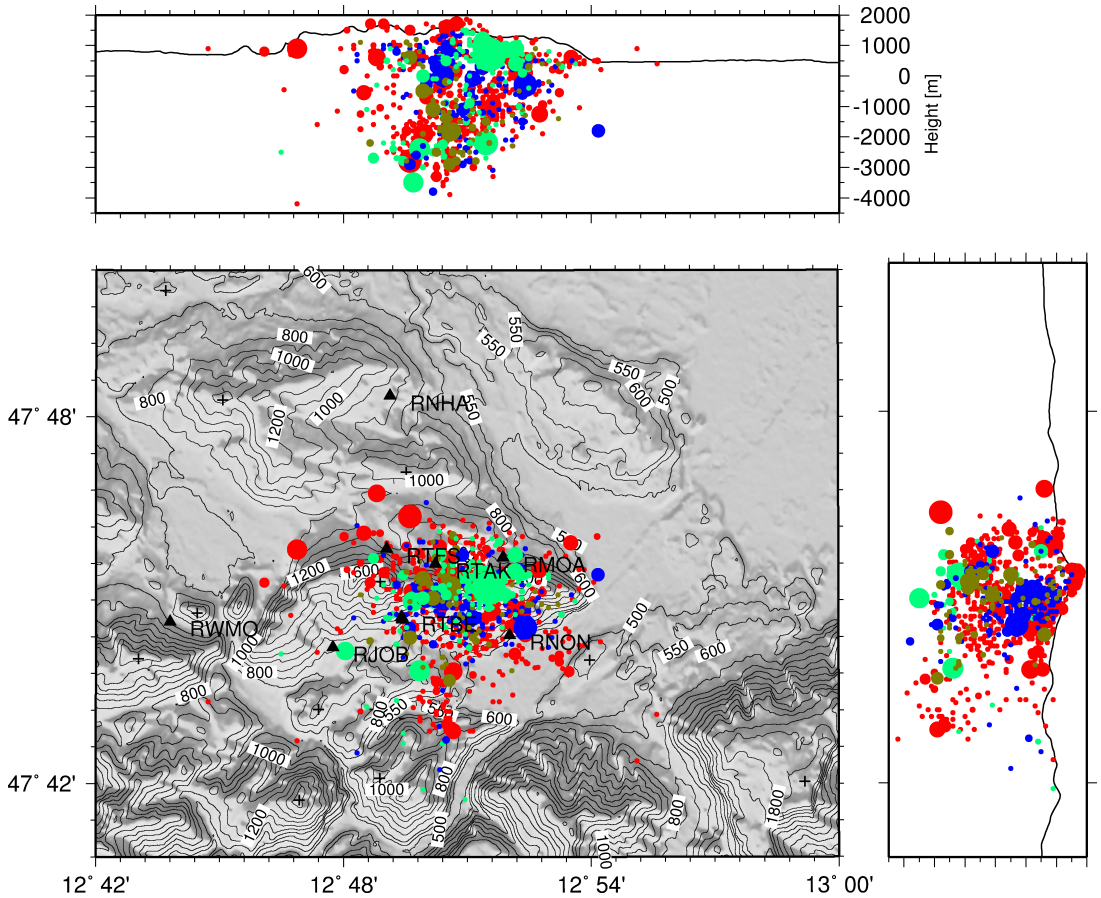
Statistics in Geophysical Sciences

- Geophysics can be subdivided by the part of the Earth studied.
- One natural division is into atmospheric science, ocean science and solid-Earth geophysics, with the solid Earth further divided into the crust, mantle and core.
- "*As mainstream physics has moved to study smaller objects and more distant ones, geophysics has moved closer to geology, and its mathematical content has become generally more dilute, with important singularities. The subject is driven largely by observation and data analysis, rather than theory, and probabilistic modeling and statistics are key to its progress.*" (see Stark, P. B. (1996))

Statistics in Geophysical Sciences: Example

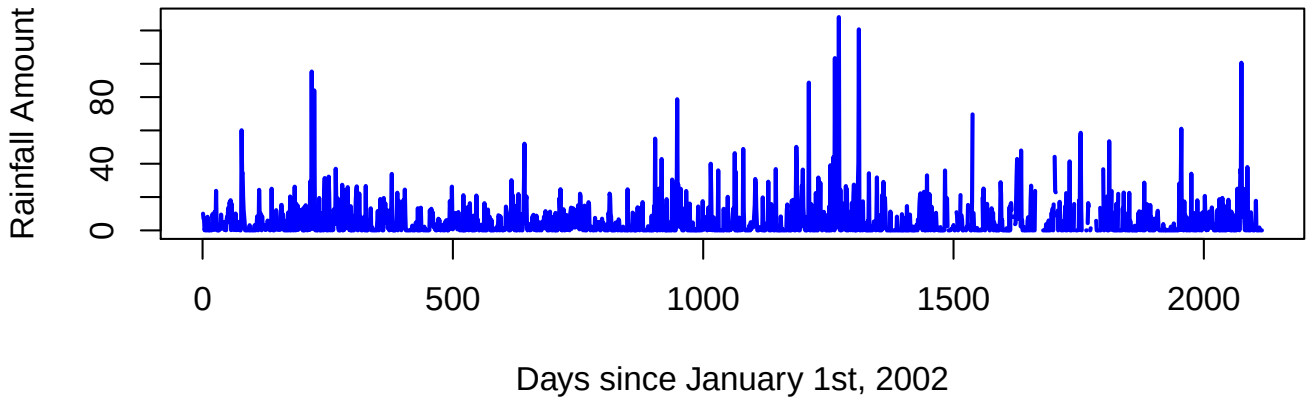
Kraft, T., Wassermann, J., Schmedes, E., Igel, H. (2006): Meteorological triggering of earthquake swarms at Mt. Hochstaufen, SE-Germany, *Tectonophysics*, Vol. 424 No. 3-4, pp. 245-258.
`http://www.geophysik.uni-muenchen.de/~igel/PDF/kraftetal_tecto_2006.pdf`

Statistics in Geophysical Sciences: Example

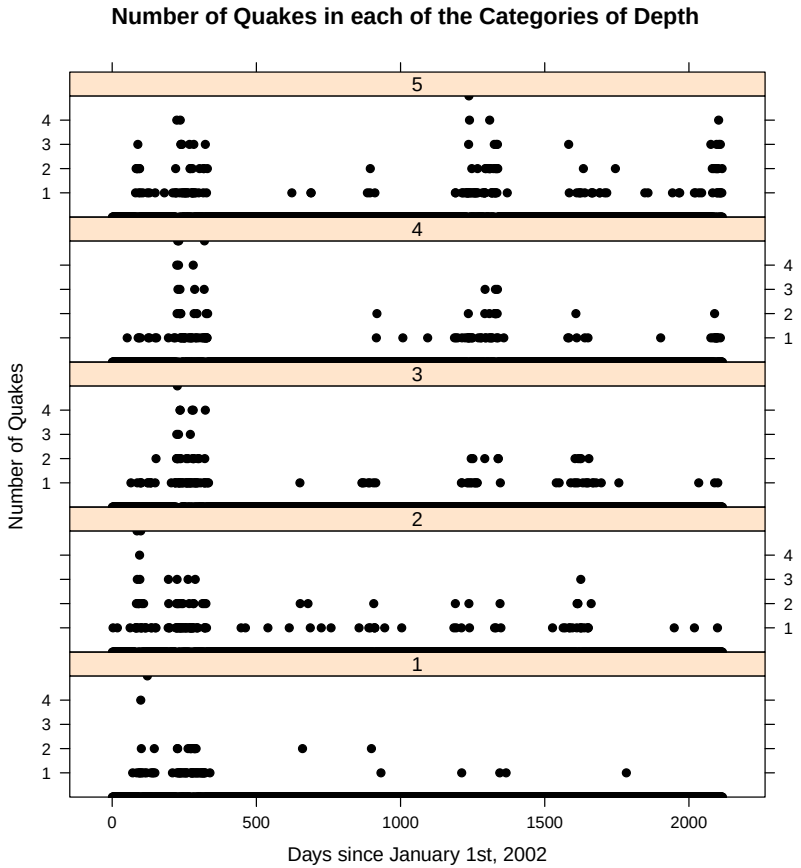


Mount Hochstaufen earthquakes

Statistics in Geophysical Sciences: Example



Statistics in Geophysical Sciences: Example



Course outline

- Probability Theory
- Descriptive Statistics
- Inferential Statistics
- Linear Regression
- Generalized Linear Regression
- Multivariate Methods

Uncertainty in Geophysics

- Our uncertainty about almost any system is of different degrees in different instances.
- For example, you cannot be completely certain
  - whether or not rain will occur at hour home tomorrow, or
  - whether the average temperature next month will be greater or less than the average temperature this month.
- We are faced with the problem of expressing degrees of uncertainty.
- It is preferable to express uncertainty quantitatively. This is done using numbers called probabilities.

Sample space

- The set, Ω, of all possible outcomes of a particular experiment is called the sample space for the experiment.

- If the experiment consists of tossing a coin with outcomes head (H) or tail (T), then

$$\Omega = \{\mathrm{H}, \mathrm{T}\} \ .$$

- Consider an experiment where the observation is reaction time to a certain stimulus. Here,

$$\Omega = (0, \infty) \ .$$

- Sample spaces can be either countable or uncountable.

Event

- An event is any collection of possible outcomes of an experiment, that is, any subset of Ω (including Ω itself).
- An event can be either:
  1. a compound event (can be decomposed into two or more (sub)events), or
  2. an elementary event.
- Let A be an event, a subset of Ω. The event A occurs if the outcome of the experiment is in the set A.
- We define

$$A \subset B \Leftrightarrow x \in A \Rightarrow x \in B \ ,$$
$$A = B \Leftrightarrow A \subset B \text{ and } B \subset A \ .$$

Set operations

Given any two events A and B we define the following operations:

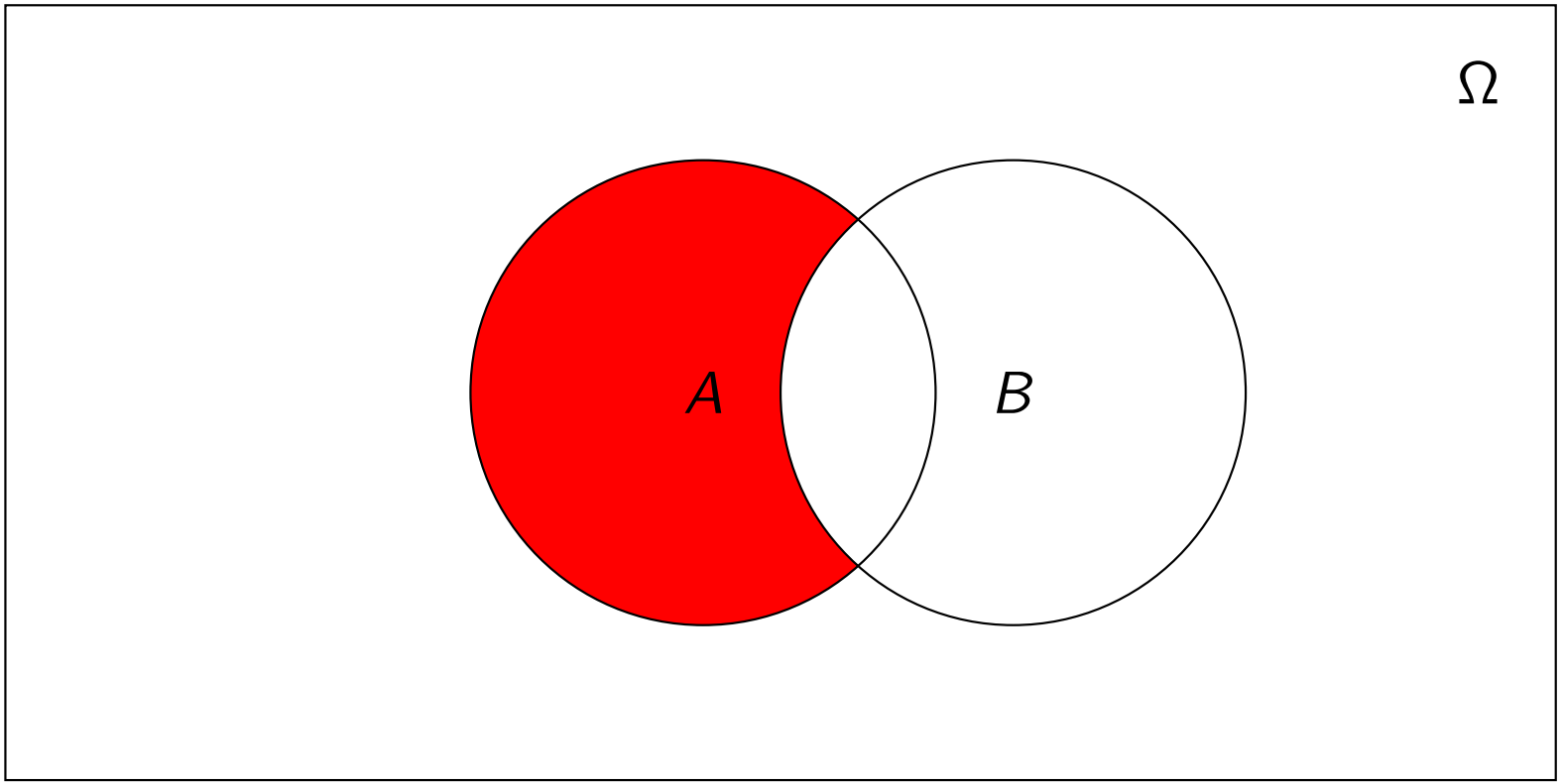
Union: The union of A and B, written $A \cup B$ is
$A \cup B = \{x : x \in A \text{ or } x \in B\}$.

Intersection: The intersection of A and B, written $A \cap B$ is
$A \cap B = \{x : x \in A \text{ and } x \in B\}$.

Complementation: The complement of A, written $\overline{A}$ (or A^c), is
$\overline{A} = \{x : x \notin A\}$.

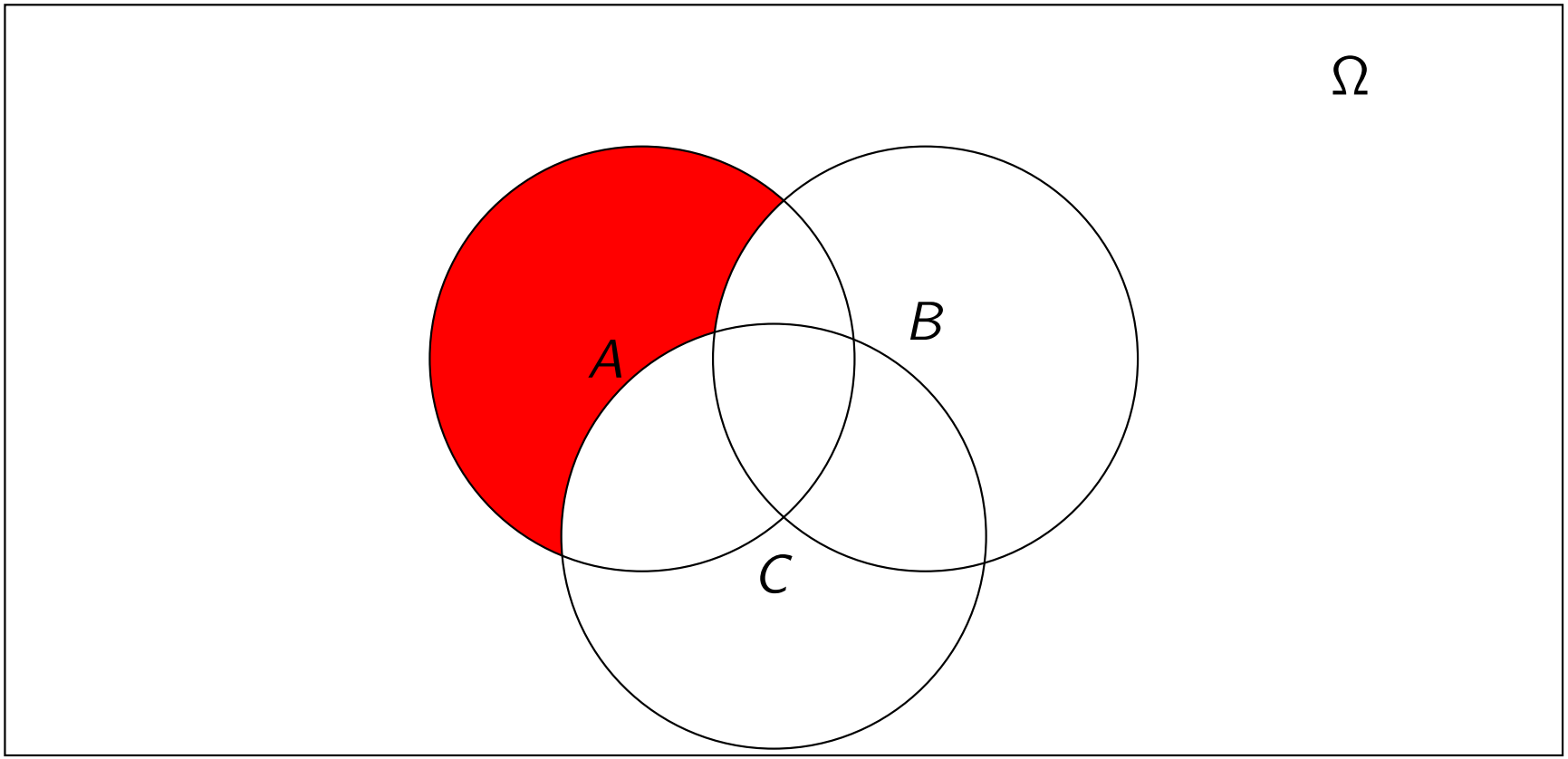
Venn diagrams

$$A \cap \overline{B}$$



Venn diagrams

$$(A \cup B) \cap (\overline{B \cup C})$$



Set operations: Example

- Selecting a card at random from a standard desk and noting its suit: clubs (C), diamonds (D), hearts (H) and spades (S).
- The sample space is $\Omega = \{\text{C,D,H,S}\}$.
- Some possible events are $A = \{\text{C,D}\}$ and $B = \{\text{D,H,S}\}$.
- From these events we can form $A \cup B = \{\text{C,D,H,S}\}$, $A \cap B = \{\text{D}\}$ and $\overline{A} = \{\text{H,S}\}$.
- Notice that $A \cup B = \Omega$ and $\overline{A \cup B} = \emptyset$, where $\emptyset$ denotes the empty set.

Properties of set operations

For any three events, A, B and C, defined on the sample space Ω,

Commutativity:	$A \cup B = B \cup A$,
	$A \cap B = B \cap A$;
Associativity;	$A \cup (B \cup C) = (A \cup B) \cup C$,
	$A \cap (B \cap C) = (A \cap B) \cap C$;
Distributive laws:	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
De Morgan's laws:	$\overline{A \cup B} = \overline{A} \cap \overline{B}$,
	$\overline{A \cap B} = \overline{A} \cup \overline{B}$.

Partition of the sample space

- Two events A and B are disjoint (or mutually exclusive) if $A \cap B = \emptyset$. The events $A_1, A_2, \ldots$ are pairwise disjoint if $A_i \cap A_j = \emptyset$ for all $i \neq j$.

- If $A_1, A_2, \ldots$ are pairwise disjoint and $\bigcup_{i=1}^{\infty} = \Omega$, then the collection $A_1, A_2, \ldots$ forms a partition of Ω.

Definition of Laplace

Théorie Analytique des Probabilités (1812)

"*The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this.*"

Definition by Laplace

For an event $A \subset \Omega$, the probability of A, $\mathrm{P}(A)$, is defined as

$$\mathrm{P}(A) := \frac{|A|}{|\Omega|} \ ,$$

where $|A|$ denotes the cardinality of the set A.

Frequency interpretation (von Mises)

The probability of an event is exactly its long-run relative frequency:

$$\mathrm{P}(A) = \lim_{n \to \infty} \frac{a_n}{n} \quad ,$$

where a_n is the number of occurrences and n is the number of opportunities for the event A to occur.

Subjective interpretation (De Finetti)

- Employing the Frequency view of probability requires a long series of identical trials.
- The subjective interpretation is that probability represents the degree of belief of a particular individual about the occurrence of an uncertain event.

Kolmogorov axioms

A collection of subsets of Ω is a sigma algebra (or field) $\mathcal{F}$, if $\emptyset \in \mathcal{F}$ and if $\mathcal{F}$ is closed under complementation and union.

Given a sample space Ω and an associated sigma algebra $\mathcal{F}$, a probability function is a function P with domain $\mathcal{F}$ that satisfies

A1 $\mathrm{P}(A) \geq 0$ for all $A \in \mathcal{F}$.

A2 $\mathrm{P}(\Omega) = 1$.

A3 if $A_1, A_2, \ldots \in \mathcal{F}$ are pairwise disjoint, then $\mathrm{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathrm{P}(A_i)$.

The calculus of probabilities

If P is a probability function and A is any set in $\mathcal{F}$, then

- $\mathrm{P}(\emptyset) = 0$;
- $\mathrm{P}(A) \leq 1$;
- $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$.

If P is a probability function and A and B are any sets in $\mathcal{F}$, then

- $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$;
- If $A \subset B$, then $\mathrm{P}(A) \leq \mathrm{P}(B)$.

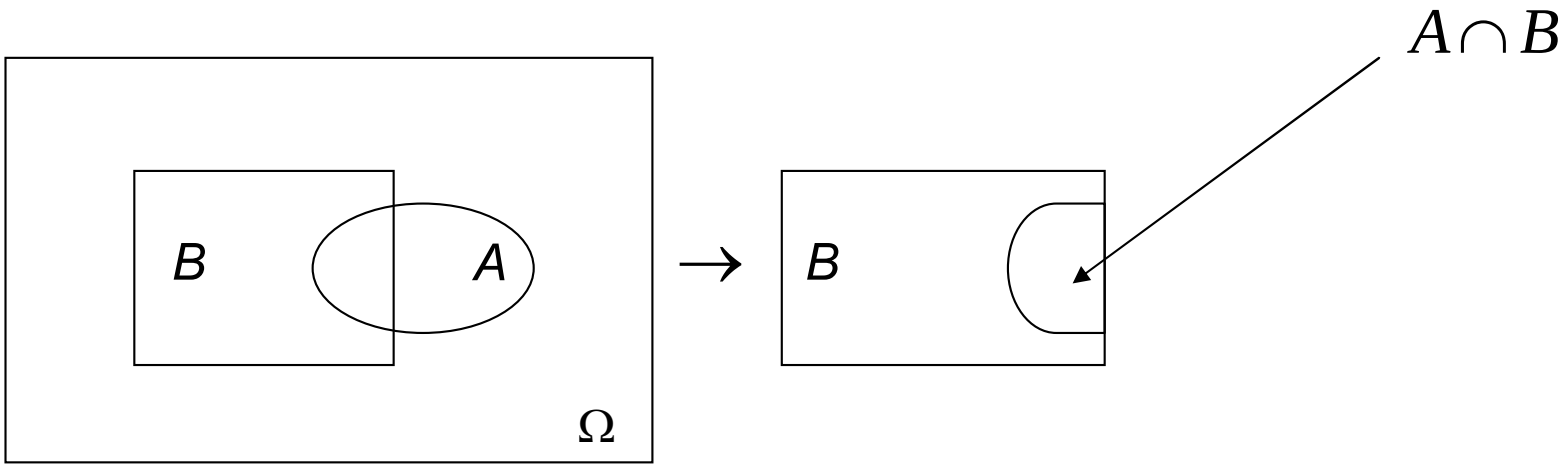
Conditional probability

If A and B are events in Ω, and $\mathrm{P}(B) > 0$, then the conditional probability of A given B, written $\mathrm{P}(A|B)$, is

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \quad ,$$

where $\mathrm{P}(A \cap B)$ is the joint probability of A and B.

Conditional probability P($A|B$)



Conditional probability: Example

	Gender		
Died from CHD	Male	Female	Total
Yes	64 473	53 003	117 476
No	223 859	265 460	489 319
Total	288 332	318 463	606 795

Table: UK deaths in 2002 from coronary heart disease (CHD) by gender

X: gender of person who died.
Y: whether or not a person died from CHD.

$\mathrm{P}(Y = \text{yes} | X = \text{male}) = ?$

Calculating a conditional probability

$$\mathrm{P}(X = \text{male}) = \frac{\text{number of male deaths}}{\text{total number of deaths}} = \frac{288332}{606795} = 0.4752 \ .$$

$$\begin{aligned} \mathrm{P}(Y = \text{yes and } X = \text{male}) &= \frac{\text{number of men who died from CHD}}{\text{total number of deaths}} \\ &= \frac{64473}{606795} = 0.1063 \ . \end{aligned}$$

$$\begin{aligned} \mathrm{P}(Y = \text{yes} \,|X = \text{male}) &= \frac{\mathrm{P}(Y = \text{yes and } X = \text{male})}{\mathrm{P}(X = \text{male})} \\ &= \frac{0.1063}{0.4752} = 0.2237 \ . \end{aligned}$$

Multiplicative law of probability and independence

Rearranging the definition of conditional probability yields:

$$\begin{aligned} \mathrm{P}(A \cap B) &= \mathrm{P}(A|B)\mathrm{P}(B) \\ &= \mathrm{P}(B|A)\mathrm{P}(A) \ . \end{aligned}$$

Two events, A and B, are statistically independent if

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B) \ .$$

Independence between A and B implies

$$\mathrm{P}(A|B) = \mathrm{P}(A) \quad \text{and} \quad \mathrm{P}(B|A) = \mathrm{P}(B) \ .$$

Law of total probability

We use conditional probabilities to simplify the calculation of $\mathrm{P}(B)$.

$$\mathrm{P}(B) = \mathrm{P}(B \cap A) + \mathrm{P}(B \cap \overline{A}).$$

Using the multiplicative law of probability, this becomes

$$\mathrm{P}(B) = \mathrm{P}(B|A)\mathrm{P}(A) + \mathrm{P}(B|\overline{A})\mathrm{P}(\overline{A}) \ .$$

In general:
If $\Omega = \bigcup_{i=1}^{\infty} A_i$ and $A_i \cap A_j = \emptyset$, then

$$\mathrm{P}(B) = \sum_{i=1}^{\infty} \mathrm{P}(B|A_i)\,\mathrm{P}(A_i) \ .$$

Bayes' theorem


Thomas Bayes (1702–1761)

Bayes' theorem is a combination of the multiplicative law and the law of total probability:

$$\mathrm{P}(A_i|B) = \frac{\mathrm{P}(A_i \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(B|A_i)\mathrm{P}(A_i)}{\sum_{j=1}^{\infty} \mathrm{P}(B|A_i)\,\mathrm{P}(A_i)} \quad .$$

Bayes' theorem

For two events A and B, provided that $P(B) > 0$,

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(B|A)\mathrm{P}(A)}{\mathrm{P}(B)} \quad ,$$

where

$$\mathrm{P}(B) = \mathrm{P}(B|A)\,\mathrm{P}(A) + \mathrm{P}(B|\overline{A})\,\mathrm{P}(\overline{A}) \quad .$$

$\mathrm{P}(A)$: prior probability
$\mathrm{P}(A|B)$: posterior probability

Bayes' theorem: Example

We use Bayes' theorem to obtain an estimate of the probability that a person who is known to have died from CHD is male:

$$\mathrm{P}(X = \text{male}|\, Y = \text{yes}) = \frac{\mathrm{P}(Y = \text{yes}|X = \text{male})\mathrm{P}(X = \text{male})}{\mathrm{P}(Y = \text{yes})} \ .$$

We obtain

$$\mathrm{P}(X = \text{male}|\, Y = \text{yes}) = \frac{0.2237 \times 0.4752}{0.1936} = 0.5491 \ ,$$

where

$$\begin{aligned} \mathrm{P}(Y = \text{yes}) &= \mathrm{P}(Y = \text{yes}|X = \text{male})\mathrm{P}(X = \text{male}) \\ &\quad + \mathrm{P}(Y = \text{yes}|X = \text{female})\mathrm{P}(X = \text{female}) \\ &= 0.2237 \times 0.4752 + 0.1664 \times 0.5248 = 0.1936 \ . \end{aligned}$$