Why

How should we modify probabilities, given that we know some aspect of the outcome (i.e., that some event has occurred)?

Definition

Suppose Ω is a set of outcomes and $\mathbf{P} : \mathcal{P}(\Omega) \to \mathbf{R}$ is a finite probability measure. Suppose $A, B \subset \Omega$ with $\mathbf{P}(B) \neq 0$. The *conditional probability* of A *given* B is the ratio of the probability of $A \cap B$ to the probability of B.

Notation

In a slightly slippery but universally standard notation, we denote the conditional probability of A given B by $\mathbf{P}(A \mid B)$. In other words, we define

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)},$$

for all $A, B \subset \Omega$, whenever $\mathbf{P}(B) \neq 0$.

For example, we can express the law of total probability (see Event Probabilities) as

$$\mathbf{P}(B) = \textstyle\sum_{i=1}^{n} \mathbf{P}(A_i)\mathbf{P}(B \mid A_i),$$

where $A_1, \ldots, A_n$ partition Ω and $B \subset \Omega$ with $\mathbf{P}(B) > 0$.

Conditional probability measure

It happens that $\mathbf{P}(\cdot \mid B)$ is itself a probability measure on Ω. We therefore refer to $\mathbf{P}(\cdot \mid B)$ as a *conditional probability measure*.

To see this, suppose $B \subset \Omega$ and $\mathbf{P}(B) > 0$. Then (i) $\mathbf{P}(A \mid B) \geq 0$ for all $A \subset \Omega$, since $\mathbf{P}(A \cap B) \geq 0$. Moreover, (ii)

$$\mathbf{P}(\Omega \mid B) = \mathbf{P}(\Omega \cap B)/\mathbf{P}(B) = \mathbf{P}(B)/\mathbf{P}(B) = 1$$

Similarly, $\mathbf{P}(\varnothing \mid B) = \mathbf{P}(\varnothing \cap B)/\mathbf{P}(B) = 0/\mathbf{P}(B) = 0$.

Finally, (iii) if $A \cap C = \varnothing$, then

$$\begin{aligned}
\mathbf{P}(A \cap C \mid B) &= \mathbf{P}((A \cap C) \cap B)/\mathbf{P}(B) \\
&= \mathbf{P}((A \cap B) \cap (C \cap B)/\mathbf{P}(B) \\
&\overset{(a)}{=} (\mathbf{P}(A \cap B) + \mathbf{P}(C \cap B))/\mathbf{P}(B) \\
&= \mathbf{P}(A \cap B)/\mathbf{P}(B) + P(C \cap B)/\mathbf{P}(B) \\
&= \mathbf{P}(A \mid B) + P(C \mid B).
\end{aligned}$$

where (a) follows since $A \cap B$ and $C \cap B$ are disjoint.

Induced conditional distribution

Therefore, we expect there to also correspond a new distribution on the set of outcomes. For $\mathbf{P}_p$, define $q : \Omega \to \mathbf{R}$ by

$$q(\omega) = \begin{cases} \frac{p(\omega)}{\mathbf{P}(B)} & \text{if } \omega \in B \\ 0 & \text{otherwise.} \end{cases}$$

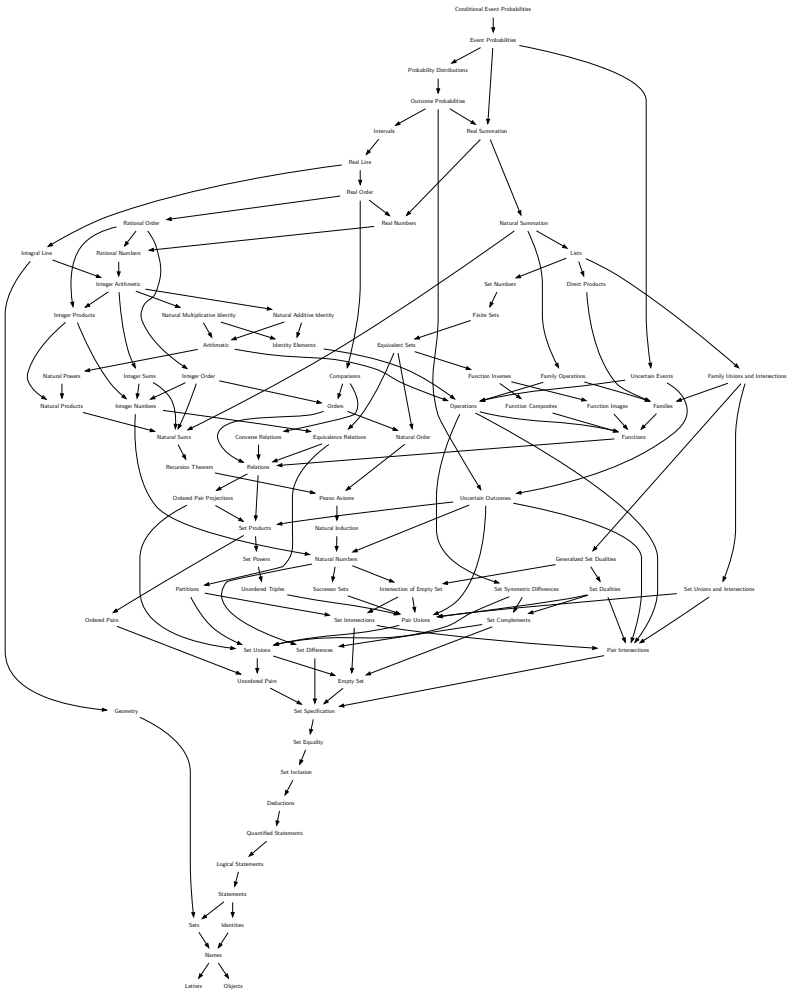
In this case $\mathbf{P}_q(A) = \mathbf{P}_p(A \mid B)$. We call q the *conditional distribution* induced by *conditioning on* the event B.

Finite intersections

As a simple repeated application of our definition, suppose $A_1, \ldots, A_n \subset \Omega$. Then

$$\mathbf{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2 \mid A_1) \cdots \mathbf{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

Many authors call this the *chain rule*. The order of the A_i is inconsequential.


Conditional Event Probabilities
Event Probabilities
Probability Distributions
Outcome Probabilities
Intervals
Real Summation
Real Line
Real Order
Rational Order
Real Numbers
Natural Summation
Integral Line
Rational Numbers
Lists
Integer Arithmetic
Set Numbers
Direct Products
Integer Products
Natural Multiplicative Identity
Natural Additive Identity
Finite Sets
Identity Elements
Arithmetic
Equivalent Sets
Natural Powers
Integer Sums
Integer Order
Comparisons
Function Inverses
Family Operations
Uncertain Events
Family Unions and Intersections
Natural Products
Integer Numbers
Orders
Operations
Function Composites
Function Images
Families
Natural Sums
Convene Relations
Equivalence Relations
Natural Order
Functions
Recursion Theorem
Relations
Ordered Pair Projections
Peano Axioms
Uncertain Outcomes
Set Products
Natural Induction
Set Powers
Natural Numbers
Generalized Set Dualities
Partitions
Unordered Triples
Successor Sets
Intersection of Empty Set
Set Symmetric Differences
Set Dualities
Set Unions and Intersections
Ordered Pairs
Set Intersections
Pair Unions
Set Complements
Set Unions
Set Differences
Pair Intersections
Unordered Pairs
Empty Set
Geometry
Set Specification
Set Equality
Set Inclusion
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Norms
Letters
Objects