# Probability Logic

Axioms and properties of probability logic: See powepoint.

## Discrete Variables

Let $X$ be an uncertain variable that can take discrete values $x_1,\ldots,x_n$. The following notation

$$p(x_i \mid a) = P(X = x_i \mid a)$$

is used to denote the probability of the proposition $X = x_i$, i.e. the variable $X$ to take the value $x_i$, given the information in proposition $a$. Note that the propositions $X = x_1$, $X = x_2$, …, $X = x_n$ are mutually exclusive. Let also $Y$ be another uncertain discrete variable with possible values $y_1,\ldots,y_n$. It can be readily verified that the following hold true.

Marginalization Theorem

$$p(x_i \mid a) = \sum_{k=1}^{m} p(x_i, y_k \mid a)$$

Total Probability Theorem

$$p(x_i \mid a) = \sum_{k=1}^{m} p(x_i \mid y_k, a)\, p(y_k \mid a)$$

Bayes Theorem

$$p(y_k \mid x_i, a) = \frac{p(x_i \mid y_k, a)\, p(y_k \mid a)}{\sum_{k=1}^{m} p(x_i \mid y_k, a)\, p(y_k \mid a)}$$

## Continuous Variables

Let $X$ be an uncertain variable that can take values on a continuous domain $x \in [x_{start}, x_{end}]$. The following notation

$$P(X \le x \mid a) \equiv F(x \mid a)$$

is used to denote the probability of the proposition $X \le x$, i.e. the variable $X$ to take value less than $x$, given the information in proposition $a$. It is referred as the cumulative probability distribution of a variable $X$. Define the probability distribution function $f(x)$ of a variable $X$ from the expression

$$P(x < X \le x + dx \mid a) = f(x \mid a)dx \tag{1}$$

It can be readily derived that

$$f(x \mid a) = \frac{dF(x \mid a)}{dx} \tag{2}$$

using the fact that the statement $X \le x + dx \mid a$ is the sum of the statement $X \le x \mid a$ and $x < X \le x + dx \mid a$ and that the statements $X \le x \mid a$ and $x < X \le x + dx \mid a$ are mutually exclusive so that using the sum rule

$$P(X \le x + dx \mid a) = P(X \le x \text{ or } x < X \le x + dx \mid a) = P(X \le x \mid a) + P(x < X \le x + dx \mid a)$$

which results in

$$P(x < X \le x + dx \mid a) = P(X \le x + dx \mid a) - P(X \le x \mid a) = F(x + dx \mid a) - F(x \mid a)$$

Using (1), one derives that $F(x + dx) - F(x) = f(x \mid a)dx$ which results in (2).

Finally, it can be readily shown that for the probability distribution function $f(x \mid a)$ of a continuous variable $X$, the following hold true.

Marginalization Theorem

$$f(x \mid a) = \int f(x, y \mid a)\, dy$$

Total Probability Theorem

$$f(x \mid a) = \int f(x \mid y, a)\, f(y \mid a)\, dy$$

Bayes Theorem

$$f(y \mid x, a) = \frac{f(x \mid y, a)\, f(y \mid a)}{\int f(x \mid y, a)\, f(y \mid a)\, dy}$$

where $Y$ is another continuous variable with probability distribution $f(y \mid a)$.

## References

1. Karl-Rudolf Koch, Introduction to Bayesian Statistics, Springer-Verlag Berlin Heidelberg 2007.