Information Fusion (IFU), Summer Semester 2023

Christoph-Alexander Holst, , Jan Segermann, Research Group Discrete Systems

Excercise 3: Probability Theory

In the third exercise important concepts of probability theory are recalled and Bayesian probability is discussed. To be able to complete this exercise, you are required to have worked through the lectures up to and including L_IFU_31.

For some of the tasks in this exercise, you are required to write down equations or visualise concepts with drawings. Our suggestion is that you either use a digital sketching tool (if available and you are accustomed to it) or that you prepare your results on paper.

3.1 Recap

It is crucial to understand the basic concepts of probability theory to be able to follow the lecture properly. Rehearse the following terms, define them formally, and describe their intuitive meaning:

1. probability density function,
2. expected value and mean value (what is the difference between them?),
3. variance, standard deviation and
4. confidence interval and confidence level.
5. [Optional] For the sample variance, it is common to apply *Bessel's correction*. What is Bessel's correction and why is it applied?

Furthermore, you may refer to Cuzzolin's introduction for some insights on probability interpretation [2, pp. 1-9].

3.2 Bayesian Probability

1. There are two main interpretations of the concept of probability: *frequentism* and *Bayesian probability*. What distinguishes the frequentist and Bayesian interpretation?
2. Conditional probability (Bayes' theorem): Let C be the event of having contracted Covid-19 and C^{c} its complement. We make a few assumptions:
   - (i) the number of 106631 cases in Germany[1] represents the actual total number of cases,
   - (ii) the population of Germany is $83.2 \cdot 10^6$ [2], and

[1] `https://www.bundesregierung.de/breg-de/aktuelles/fallzahlen-coronavirus-1738210`, 04.05.2022

[2] `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/_inhalt.html`, 04.05.2022

- (iii) we use a self-test for Covid-19 with the following properties[3]: The test has a probability of $P(T \mid C) = 0.9537$ to produce a positive test result T under the condition that a person suffers from Covid-19 (true positive result). Furthermore, it has a probability of $P(T \mid C^{\mathrm{c}}) = 0.0087$ of a false positive result.

Living in Germany, what is the probability of having Covid-19 given a positive test? Are there problems with the assumptions made? Is this approach (including the assumptions) still correct if you show symptoms of Covid-19 (e.g. fever)?

3. Proof the general rule of total probability (cf. lecture L_IFU_31 slide 25).

4. One of the main drawbacks of modelling evidence in probability theory is the inability to distinguish between disbelief and lack of belief [3, p. 91][4] At first glance, this may seem unintuitive. So, we will discuss the inability to express ignorance within probability theory using the example provided in [3] (example originally given by Shafer in [4]).

   - Let A be an event (or proposition), A^{c} be its complement, and S be the universe of discourse (or event space). We define $\mathrm{Bel}(A)$ to be a function expressing our belief or evidence that A is true. In Bayesian probability $\mathrm{Bel}(A) = P(A)$. What is then $\mathrm{Bel}(A \cup A^{\mathrm{c}})$?
   - Shafer's example refers to the question, whether there are living beings around the star Sirius. The proposition A_1 corresponds to the existence of life and A_2 to the absence of life. Assuming total ignorance (we have no clue), which values would you assign $\mathrm{Bel}(A_1)$ and $\mathrm{Bel}(A_2)$?
   - A different situation: Assume you acquired a new two-sided coin. Let's assume that the coin, if tossed, cannot land on its edges, only events B_1 „head side up“ and B_2 „tail side up“ are possible. You have not tossed the coin yet, but the coin looks well-balanced. You toss the coin. What are $\mathrm{Bel}(B_1)$ and $\mathrm{Bel}(B_2)$?
   - In these examples, what is the fundamental difference in deriving the beliefs? Does it matter?

5. After having completed this exercise, which categories in the taxonomy of ignorance do you think probability theory is well suited for? For which categories is it not suited?

[3] Beijing Hotgen Biotech Co., Ltd, Coronavirus (2019-nCoV)-Antigentest

[4] The section in Salicone's book about representing ignorance in probability theory presupposes knowledge about the Dempster-Shafer-Theory. For this exercise it is not absolutely necessary to read the cited pages, but feel free to do so if you want.

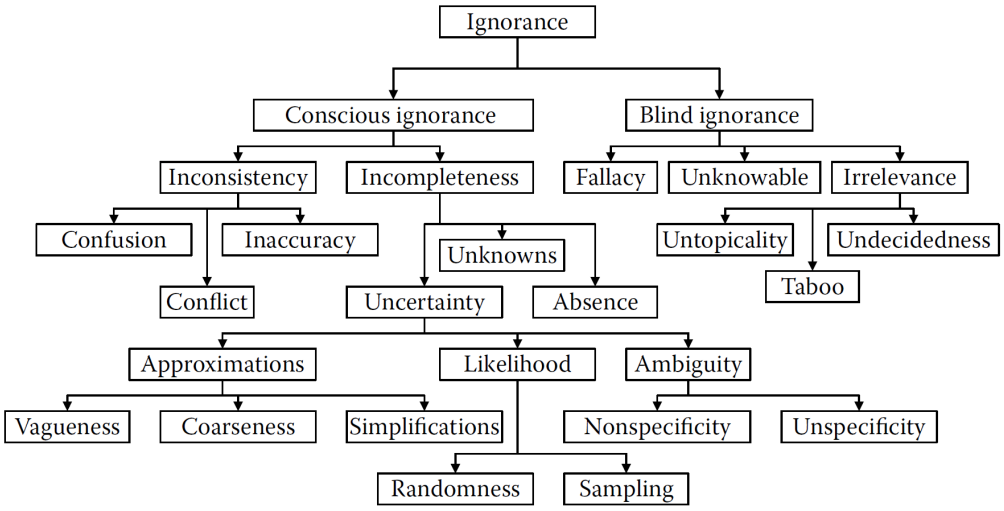

Figure 1: A taxonomy of ignorance given in [1, p. 53].

References

[1] Bilal M. Ayyub and George Jiri Klir. *Uncertainty Modeling and Analysis in Engineering and the Sciences*. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[2] Fabio Cuzzolin. *The geometry of uncertainty: the geometry of imprecise probabilities*. Springer Nature, 2020.

[3] Marco Prioli and Simona Salicone. *Measurement Uncertainty within the Theory of Evidence*. Springer Series in Measurement Science and Technology. Springer International Publishing, 2018.

[4] Glenn Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.