General Discussion

Statistical thinking is indispensable in an uncertain world. It is a requirement for every-day life as well as for the professional decision maker in domains such as law and medicine and, last but not least, it is important for scientific inference. This dissertation demonstrated how insight into crucial statistical problems can be fostered and statistical thinking can be taught.

Chapter 1 showed that for Bayesian inferences one can abandon the concepts of probabilities and percentages, which are not well-adjusted to human thinking. It was demonstrated that the natural frequency approach – well-established in simple Bayesian situations – is also of help in complex Bayesian settings. With the help of natural frequencies, relevant decision situations such as in medical diagnoses or for legal verdicts can be profitably represented, even when multiple evidence is available. Open questions for future research are:

What are the limits of the natural frequency approach? In the case of 10 cues, it might be impossible to remember all relevant frequencies in order to come up with the Bayesian solution. It is important to note that for multiple-cue situations one has to distinguish real-life decisions from text problems. In text problems, memory constraints do not play a role, since the information relevant for the decision does *not* have to be retrieved from memory but can be read from the instruction sheet. This, by the way, might also explain why in our Study 1 performance using 2 cues is the same as performance using 3 cues: In order to solve the usual Bayesian text problems *only both natural frequencies pertaining to the last cue* have to be taken into account – regardless of the amount of previous information on other cues. These frequencies fully determine the frequentistic Bayesian solution (see Figure 1.1 on the right). Thus, the prediction is that when provided with Bayesian text problems – no matter how many cues are provided – participants can compose the correct inference from the two natural frequencies mentioned above.

What if multiple cue information is not presented in the form of an instruction sheet? Gigerenzer, Todd and the ABC Group (1999) suggested fast and frugal heuristics that do not rely on all available information but still have a satisfactory performance. For instance, the "Take The Best" heuristic uses just the best cue and yet performs surprisingly well (see chapters 4, 5 in Gigerenzer, Todd and the ABC Group, 1999). The present dissertation showed that natural frequency algorithms that *integrate all available information* are still possible in complex Bayesian situations (see chapters 1 and 2), while Gigerenzer, Todd and the ABC Group (1999) propose simple heuristics that *ignore a part of the available information*. Ongoing work is melting both lines of work by fleshing out the conditions under which people are Bayesians and those under which they are fast and frugal (Martignon & Laskey 2001; Martignon and Krauss, 2001).

Future work has also to look at the pedagogical side. The natural frequency approach for simple one-cue inferences is by now well established in journals of medicine and law (Gigerenzer, Hoffrage & Ebert, 1989; Krauss & Hertwig, 2000). One goal is to spread the new results on complex Bayesian inference to these relevant realms.[35] Another field of application is teaching statistics to pupils. Currently, in the DFG-Project *Verbesserung des Stochastikunterrichts in der Schule* the contents and proposals of all 3 chapters of the present dissertation are tested in intervention programs at German secondary schools with pupils of the 12th and 13th grade (see, e.g., Jahresbericht des Johannes-Scharrer-Gymnasium Nürnberg, 2001).

One of the most intriguing and fascinating appetizers for a course in probability is the Monty Hall problem. In Chapter 2, it was shown that this probabilistic brain teaser can be deconstructed with the help of cognitive psychology. By implementing manipulations in the wording of the problem based on four elements from the cognitive psychologists' toolbox, namely the *natural frequency* concept, the *mental model* concept, *perspective change*, and the *less-is-more* effect, we could foster insight into the problem's structure. Furthermore, it was shown that these manipulations are also of help in other related problems of probability theory.

The general pedagogical approach of this chapter also deserves some remarks. We analyzed an "intuitive" solution, decomposed this solution into psychological elements, and then paved the pathway for participants towards this solution by manipulating the wording according to the discovered elements. In contrast to the common "task analysis", this suggests a new pedagogical approach: "Solution analysis" extracts the method most frequently chosen by naive participants to solve the problem correctly and analyzes this solution with regard to underlying psychological elements. Exploiting these elements through respective manipulations in the wording might then guide new participants towards

[35] By focusing strictly on the medical domain in Chapter 1, one should not forget that there are also other domains where experts could profit from the present findings: Judges, for instance, are often faced with more than one suspect, with more than one piece of evidence, or, with unclear evidence.

this "intuitive way" of solving the problem. Future research has to reveal the potential of this approach for other domains.

In Chapter 3, finally, we demonstrated that misconceptions about the meaning of significance are still wide-spread – even among statistics teachers. This dissertation suggested a pedagogical procedure that guarantees a well-informed handling of NHST, without misconceptions nor beliefs in myths. For instance, to avoid the common belief that significance can say anything about the probability of hypotheses, we proposed to present the approach that can actually do this, namely Bayesian statistics. Expressing not only Bayes' rule but also significance testing in terms of conditional probabilities delivers a clear distinction about the underlying principles of both approaches.

Statistics is the subject most taught in German universities. In approximately 50% of the departments of German universities statistics is taught, and we estimate that every third student during his or her studies is faced with statistics courses (Gigerenzer & Krauss, 2001, p. 1). Since significance testing is usually one of the main contents of these courses, the understanding of a significant test result should be a matter of course rather than a fluke.