The Mathematics Enthusiast

Volume 12
Number 1 *Numbers 1, 2, & 3*

Article 15

6-2015

A Cognitive Framework for Normative Reasoning under Uncertainty, and Reasoning about Risk, and Implications for Educational Practice

Sylvia Kuzmak

A Cognitive Framework for Normative Reasoning under Uncertainty, and Reasoning about Risk, and Implications for Educational Practice

Sylvia Kuzmak

Rise Coaching and Consulting LLC NJ, USA

Abstract: Clarifying what is normative or appropriate reasoning under various circumstances provides a valuable reference for guiding what should be taught, and, in contrast, what should not be. This paper proposes a cognitive framework for viewing normative reasoning and behavior under uncertainty, including the applying of knowledge of probability and statistics in real world situations; and identifies implications for educational practice. Factors relevant to normative reasoning under uncertainty that are addressed within the framework include: risk of misapplying statistics knowledge, involvement of mathematical and non-mathematical reasoning, knowledge of real world domains and situation/application detail, and existence of expert consensus. The cognitive framework is illustrated using examples of reasoning about risk, including industry standards for risk management. The work of Kahneman and Tversky, G. Gigerenzer, and others is related to and contrasted to the framework presented.

Keywords: reasoning under uncertainty, statistical reasoning, probabilistic reasoning, risk, improving probability education, improving statistics education, misapplication of statistics, statistics application.

Introduction

Clarifying what is normative or appropriate reasoning under various circumstances provides a valuable reference for guiding what should be taught, and, in contrast, what should not be. This paper proposes a cognitive framework for viewing normative reasoning and behavior under uncertainty, including the applying of knowledge of probability and statistics in real world situations; and identifies implications for educational practice. In sections below, factors relevant to normative reasoning under uncertainty are identified, illustrated with examples, and related to the research literature on reasoning under uncertainty. In particular, examples involving reasoning about risk are addressed, including reasoning reflected in industry standards for risk management. In the final sections, the factors are integrated into a cognitive framework; and implications for educational practice are identified.

In real world situations, we are often in the position that the outcome of a situation, which is subject to uncertainty, matters to us. To reason and behave normatively at such times is important, since doing so helps to bring on potential benefits and/or to stave off potential difficulties. Such real world situations draw interest and are engaging, and typically call for action, because the results matter. Such real world situations also are aptly described as involving risk. Not only is the outcome of the situation uncertain, with alternative possible outcomes, but the possible outcomes have positive or negative impact, so that there is risk that a positive outcome will not occur, and/or risk that a negative outcome will occur. By addressing what is normative reasoning and behavior in such situations, the cognitive framework presented here applies in general to reasoning about risk.

As a simple example, consider observing the rolling of a pair of six-sided dice. There is uncertainty in the outcomes, but unless the rolling occurs in the context of a game or other real world consequences, the outcomes don't really matter. Now, consider that the rolling of the dice is occurring in the context of gambling, and that you are about to place a large sum of your money as a bet on the

The Mathematics Enthusiast, **ISSN 1551-3440, vol. 12, no. 1,2&3**, pp. 140-156

outcome of the roll. Now the outcome is more important, the situation is more engaging, you are interested in your options for action that may make a difference in the situation, and there is risk. As another example, consider that you are a young person and occasionally contemplate your own mortality, but realize that, due to your generally safe environment and healthy habits, your odds are good that you will live a long life; and so the issue of your possible early death is not of real concern to you. Now consider that you and your spouse are just starting a family. Although the probability that you will die relatively young is still low, now the possibility of your early death is of concern to you, since it would have a great financial and otherwise life-impacting effect on your remaining family. The situation regarding your mortality is now more important, more engaging, involves risk, and has led you to consider possible actions, including buying life insurance.

The focus of the framework presented here is not just on people's judgments of probability of outcomes in situations involving uncertainty, but a broader sense of reasoning under uncertainty that includes consideration of risk, perceived consequences of outcomes, and human actions/ behavior in that context.

Mathematical and Non-mathematical Reasoning

Historically, the research area of "reasoning under uncertainty" in cognitive psychology and decision science has been closely identified with the mathematics of probability and statistics. In a seminal paper in the field, Tversky and Kahneman (1974) reported their research in which adults had been posed written problems calling for them to reason and make judgments under uncertainty; and the authors concluded that their subjects showed "biases" and "errors" in their judgments, using reasoning heuristics, such as *representativeness* and *availability*, while not being influenced by relevant mathematical information provided in the problems, such as prior probabilities or base rates, and sample sizes. The authors describe the representativeness heuristic as judging the probability that an object A belongs to a class B, or that an event A originates from a process B, "by the degree to which A is representative of B, that is, by the degree to which A resembles B" (p.1124). An example provided is the case of being given a written personality sketch of a person and being asked to judge the probability that the person has a particular occupation, such as librarian, engineer, or lawyer; and making the judgment based on the person's similarity to one's stereotype for the occupation. The availability heuristic refers to assessing "the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind" (p. 1127). An example provided is the case of "assess[ing] the risk of heart attack among middle-aged people by recalling such occurrences among one's acquaintances" (p.1127). The body of research on heuristics and biases in reasoning and judgment has grown over the decades (for example, see compilations of research in Kahneman, Slovic, & Tversky, 1982; and Gilovich, Griffin, & Kahneman, 2002). A continuing theme within the body of research has been to note the presence of error in human judgment under uncertainty with respect to not understanding basic principles of probability and statistics and/or not applying them when appropriate. In a review of his work and others' on heuristics and biases, Kahneman (2011) noted that a takeaway of the work is "that there are distinctive patterns in the errors people make" in their "judgments and choices" (p.3).

Others have objected to the negative view of human ability and performance that has come along with the "heuristics and biases" literature, and have observed that people, in their daily activities do regularly encounter uncertainties, and show competence in dealing with them, although they may use heuristic reasoning, and not be applying a knowledge of probability and statistics in these situations. For example, Cohen (1981) noted regarding research on heuristics and biases, that subjects' performance on unfamiliar tasks and in laboratory conditions, is not an adequate basis for judging their rationality or competence; and that, in some cases, subjects may legitimately interpret the problem differently than does the experimenter, and respond rationally within their interpretation. Also, he noted that some of the research is merely a "test of intelligence or education" (p.325), demonstrating "a lack of mathematical or scientific expertise" (p. 325) that is unreasonable to expect that all would have. Kahneman (2011) has defended the negative focus of the heuristics and biases literature by drawing an analogy to the legitimate attention to the study of disease in the field of medicine (p.4).

To call out "error" in judgments and choices, as is done in the heuristics and biases literature, is to presume some notion of what is normative or appropriate to do under the circumstances in which the error was observed. Cohen (1981) criticized researchers' tendency to assume an "inappropriate normative theory" (p.328) drawn from logic and statistics, not recognizing that many important normative issues are still controversial among experts in the field. Lopes and Oden (1991) criticized the heuristics and biases literature for its "normative models drawn from probability theory, economics, and logic" (p.201), inadequate to serve for the broad range of human decision making under uncertainty. They noted that, in the field of artificial intelligence, where intelligent performance of machines is the goal, heuristic reasoning is viewed positively and researchers "seek … out and embrace" (p.209) the use of heuristics. Lopes and Oden advocate the positive perspective that observed human performance "signal[s] the operation of a quite different kind of intelligence than is implied by conventional notions of rationality, an intelligence reflecting the properties of the [human's] massively parallel computational system that has evolved to meet the requirements of existence in a noisy, uncertain, and unstable world" (p.201).

Gigerenzer (1996) has critiqued the heuristics and biases literature as having too narrow a view of normative reasoning, and being too vague regarding the cognitive processes involved in heuristic reasoning. Regarding the narrow view of norms, he noted that, within the literature, there is a "practice of imposing a statistical principle as a norm without examining content" (p.593) of the problem posed; he says, "A convenient statistical principle, such as the conjunction rule or Bayes's rule, is chosen as normative, and some real-world content is filled in afterward, on the assumption that only structure matters. The content of the problem is not analyzed in building the normative model, nor are the specific assumptions people make about the situation" (p.592). So the failure of researchers to more broadly analyze the content of the problem and reasonable potential interpretations, then bypasses analysis of information that is needed to determine what is normative reasoning under the circumstances. As an example, Gigerenzer noted that the words "probable" and "and" used in a written problem, may legitimately be interpreted with natural language meanings by subjects, while researchers intend their interpretation as mathematical or logical terms; and the subjects' reasonable interpretations bear on what is normative reasoning under the circumstances.

Instead of viewing heuristic reasoning as being subject to error due to ignoring some information, Gigerenzer and Gaissmaier (2011) make the point that the use of heuristic reasoning that "ignor[es] part of the information can lead to more accurate judgments than weighting and adding all information, for instance for low predictability and small samples" (p.451). They note that "for many decisions, the assumptions of rational models [defined by logic or statistical models] are not met, and it is an empirical rather than an a priori issue how well cognitive heuristics function in an uncertain world" (p.451).

Regarding normative reasoning under uncertainty, then, a conclusion here is that both *mathematical reasoning* (using mathematical logic, probability, and statistics) and *non-mathematical reasoning* (such as heuristic reasoning and language interpretation) have a role. Non-mathematical reasoning is used here as a broad umbrella term to contrast to mathematical reasoning, the latter referring to reasoning using the formal subject matter of mathematics, including concepts, principles, and techniques of logic, probability, and statistics. It should be noted that what is called here non-mathematical reasoning nonetheless may have quantitative aspects, such as judging degrees of similarity and frequency. It is beyond the scope of this paper to provide an inventory and elaborated description of this rich and important category that is simply bundled here as non-mathematical reasoning. For many real world problems involving uncertainty, probability and statistics are powerful tools applicable to the circumstances, and are widely acknowledged as the norm to use under those circumstances. However, there are also situations involving uncertainty in which models based on probability or statistics do not apply, e.g., when the required assumptions for the model are not met; and yet, effective reasoning leading to adaptive judgments and behavior may proceed using heuristic reasoning. For some situations, there may look like a way to apply mathematics, but that application of mathematics in the situation may not be normative. In the later section that introduces the cognitive framework for normative reasoning and behavior under uncertainty, an example is presented which illustrates the need for caution in applying mathematical reasoning and the role of non-mathematical reasoning in normative reasoning and behavior.

Risk of Misapplying Knowledge of Probability and Statistics

The power and value of knowledge of probability and statistics lies in its appropriate application in real world situations. If the applicability of the math in certain circumstances is controversial, or if one applies the math to a situation without care to attend to relevant detail, or if one's knowledge is partial or lacks the firmness to ensure appropriate application to the circumstances, then there is real potential for misapplication. In such cases, to apply such knowledge involves risk and calls for awareness of the risk.

Misapplications of knowledge in probability and statistics may also occur intentionally with self-serving motivation, what has been called "lying with statistics," done by those who aim to take advantage of an audience who is inattentive, unsuspecting, and/or weak in statistics knowledge, with the goal to reap advantage for themselves to the detriment of others. Due to this possibility, the use of reported statistical information in reasoning involves risk and calls for awareness of the risk.

The risk of misapplication is greater, the greater the likelihood that the application of knowledge is in error, and the greater the negative impact of the error(s). Regarding negative impacts, misapplications may lead to flawed statistical results, e.g., inaccurate estimates or false reports of statistical significance of results. And then those flawed results may be used as a factor in decision-making, leading to unsound decisions with follow-on negative consequences. Whether the situation relates to results in medical science and patient treatment choices, business data and product promotion decisions, political polling and decisions on allocation of campaign resources, environmental impact data and setting of government environmental policy, product information in consumer reports and personal purchasing decisions, or other matters, the negative consequences of misapplication of knowledge of probability and statistics may be great.

Regarding the likelihood of a person's application of probability and statistics in a situation being in error, relevant is the chronic finding concerning modern education in probability and statistics that the subject matters are difficult for students to learn and for teachers to effectively teach. In a frequently cited research review from over 25 years ago, Garfield and Ahlgren (1988) concluded: "… despite the enthusiastic development of new instructional materials, little seems to be known about how to teach probability and statistics effectively" (p. 45). In a more recent research review, Tishkovskaya and Lancaster (2012) conclude that "Despite the widespread emphasis on reform in the teaching of statistics and the increase in papers on statistics education in the research literature, statistics is still viewed as a discipline with a need for significant improvement in how students are educated (Garfield & Ben-Zvi, 2008)" (p. 2).

Considering the circumstances just reviewed, including widespread weak knowledge of probability and statistics, yielding potential for misapplication of probability and statistics, and the range of negative impacts that may result from misapplication; therefore, in present circumstances, it is reasonable to view the risk of misapplying probability and statistics as significant.

There is evidence that even those who have taken courses in probability and statistics as part of their professional training, and who use statistical methods in their profession, fail to apply those methods appropriately, to the detriment of their professional work (Simmons, Nelson, & Simonsohn, 2011; Ioannidis, 2005; John, Loewenstein, & Prelec, 2012). This is presumably due at least in part to problems in learning the subject, and may also occur intentionally with self-serving motivation, e.g., manipulating data to be able to report statistically significant results and get one's work published. It is reasonable to suppose that the follow-on negative consequences of such professional misapplication of probability and statistics, including wasted effort and opportunity lost, incurred across the broad range of professional fields, is great.

In this paper, I raise the point that it is right to take into account in reasoning under uncertainty the risk of misapplying knowledge of probability and statistics. In a sense, the mathematics is a double-edged sword—powerful when applicable and applied appropriately, and also presenting great risk, both from unintentional misapplication, and from intentional misuse. Regarding the unintentional misapplication, it is a virtue to appreciate the limits of one's own state of knowledge. Incompleteness and lack of firmness in one's own knowledge is a true source of uncertainty, as one analyzes a situation and faces decision-making under uncertainty. Self-awareness and accuracy of perception of one's own weaknesses in knowledge supports normative reasoning and behavior for the individual in those circumstances. Recognizing the virtue of such awareness is akin to appreciating the wise saying, "A little knowledge is a dangerous thing." Regarding the intentional misuse, given the reality of its presence (e.g., in commercial advertising, political campaigning, and business financial reporting), again, it is a virtue to consider this factor in reasoning under uncertainty, being circumspect in evaluating reported statistical results and in integrating them into one's own reasoning.

Illustrations and Discussion

To illustrate the potential for error in applying knowledge of probability and statistics, two real world examples are described below, that involve drawing inferences about a population from a sample. Discussion follows regarding the potential errors and negative consequences.

Two examples. Consider the case of conducting an opinion survey and compiling the results; or sampling products from a manufacturing process, and measuring the sample's characteristics. In the case of the survey, for each item on the survey, the percentage of people in the sample making each possible response is tabulated. In the case of the manufacturing process, for each sampled product, measurements are taken for a set of characteristics of the product, and the mean measurement (arithmetic average) for each of the characteristics (across the products in the sample) is calculated. One may now conclude that the opinion percentages in the sample tell us the opinion percentages in the population as a whole (or close to it); and one may conclude that the manufacturing process, in general, produces products with the mean measurements obtained from the sample (or close to it); but if one were to do so, one would be misapplying knowledge of probability and statistics. What are the errors?

Need evidence that the sample is representative of the population. Steps need to have been taken to ensure that the sample is representative of the population about which one wishes to make inferences. In the case of a survey, this means that the demographic characteristics of the sample should match the demographic characteristics of the population. So, for example, one should sample the full age spectrum, and not just residents at a retirement community or just newly registered voters. If one wants to predict the outcome of an election, then one should sample likely voters (ones that have a history of voting regularly), and certainly not just unlikely voters. In the case of the manufacturing product sample, one should sample from the range of products about which one wants to draw conclusions, and not just the products that are least costly to sample, or for which one expects the "best performance" for the characteristics; and one should sample from the full spectrum of manufacturing conditions, e.g., from both day and night shifts, and from a range of typical manufacturing equipment conditions. To summarize, the procedures for data collection and analysis should be designed and be faithfully executed to ensure that the sample is representative of the population. In reality, samples may be collected in an unplanned manner; or in a manner to minimize collection time and/or cost, without regard to the resultant representativeness of the sample; or in a manner designed to manipulate results to a desired end instead of focusing on having a representative sample. Large data sets may be collected using multiple personnel, each using different and undocumented procedures; and respondents in a study may be anonymous, without demographic information collected; and such circumstances limit the usefulness of the data set for making inferences about the population. Random sampling from a population is a technique to obtain a representative sample, but that technique has not necessarily been faithfully used.

The mean can be a misleading measure of central tendency. Just knowing the mean (arithmetic average) of a sample, or of the entire population of focus for that matter, does not tell one what the underlying distribution of the population is. Although one may expect the bulk of a distribution to be around the population's mean, and that may often be true, such as for a normal distribution; it is also possible for it not to be true, such as is the case for bimodal or highly skewed distributions. For example, household income in the United States is a highly skewed distribution, with a long tail extending up into the super rich range; and the population mean income is much higher than the income of the central bulk of the population. For that reason, the median and not the mean is used in government reporting of US household income trends. In the case of an opinion survey, one may have a bimodal distribution if two groups comprising the population respond very differently to a question (e.g., teenagers and adults). In the case of measuring products from a manufacturing process, one may have a bimodal distribution if settings of the manufacturing machinery shift abruptly (e.g., accidentally due to human error, or due to mechanical failure) during the time when the sample is being collected. In these examples, the mean of the sample is a misleading measure of central tendency, since it does not inform one about where the bulk of the population lies, and, indeed, it may be in a range where very little of the population falls.

Need to take into account the sample size, whether sampling is random, and population variance. Since there is variation in a population, evident in its distribution, a single observation taken randomly from the population may be from anywhere in the possible range. When a random sample is taken from any population, the larger the sample size, the sample mean approaches being normally distributed with the mean being the same as for the population from which the sample is drawn (by the classical Central Limit Theorem). Thus, in the examples, for the sample mean to be informative regarding the population mean, the sample must be random and of sufficient size; and the larger the variance of the population, the larger the sample size is needed to bound the estimate of the population mean. The general caution is not to make wide-sweeping conclusions about a population based on a few observations or a small sample, and when the sample is not random.

Need to take into account required assumptions when applying statistical techniques. Based on an opinion survey, one might want to assess whether teenagers and adults differ in their opinions on a particular issue. Or, similarly, based on manufacturing product measurements, one might want to assess whether production quality differs for the day shift and the night shift. Both these inquiries relate to whether two groups are different. There are statistical tests designed to assess whether two groups are different, and these tests have assumptions that apply for their use. A commonly used test for difference between two groups is the Student's t-test, which assumes the data from the two populations are each normally distributed with the same variance. If the assumptions for the use of a statistical test do not hold for the situation of study, then to use that statistical test in that situation is not appropriate, and to do so may yield misleading results. Data is not always normally distributed, so use of techniques that assume normally distributed data is not always appropriate. For testing whether two groups differ, there are non-parametric statistical tests that apply no matter what the two population distributions are, such as the Mann-Whitney U-test, which is based on ranking of scores. Unless the assumptions for a technique's use are met in the situation to which it is applied, the results may be misleading, and lead to unsound decisions and negative follow-on consequences.

Need to consider practical significance along with statistical significance, notably when sample size is large. Continuing the example of looking for whether there are differences between two groups, and assuming that an appropriate statistical test is applied, let us say that a statistically significant difference is reported. So, for example, for the opinion survey, teenagers and adults are found to differ statistically significantly in their support for increased federal funding for low interest student loans. For the manufacturing product measurements, product quality is found to be statistically significantly different for the day and night shifts. A natural conclusion in both these cases may be that the results are meaningful, relevant, and noteworthy. However, that may not be the case. It is not just the statistical significance of the difference that matters, but also the magnitude of the observed difference between groups. A very small difference between groups may be detected by a statistical test if the sample size is large enough. In the two examples, if the sample size is large (e.g., 200), the actual magnitude of the observed difference between the two groups (that are statistically significantly different) may be quite small, not large enough to warrant treating the two groups differently, or warranting any action based on the observed difference between groups. An example of an abuse related to this principle is when a researcher seeks to conduct a study that will yield statistically significant results so that it will be publishable, and designs a study for which there is low effort and cost to include additional subjects or items in the study, accumulates large sample sizes for the two groups, and obtains results that show a statistically significant difference between groups, and show only a small observed magnitude difference between the groups, thus producing results that are of limited usefulness or merit.

The above real world examples and discussion provide a brief view into specific potential misapplications of knowledge of probability and statistics and potential follow-on negative impacts, to illustrate that the risk of misapplication of knowledge is broadly present and practically significant in real world situations. It is a true source of uncertainty, of which it is a virtue to be aware and to address with reason; and, it is right to be included in a description of normative reasoning under uncertainty.

Real World Domain Knowledge and Situation/Application Detail

Reasoning under uncertainty involving the application of knowledge of probability and statistics involves reasoning about the real world situation to which the math is applied. The illustrations and discussion in the previous section illustrate that mathematical concepts, principles, and techniques are mapped onto a real world situation, and reasoning proceeds within the domain of the situation, referencing relevant details of the situation and the mathematics that applies. For example, the concept of random selection must be applied within the application detail of selecting and procuring people to participate in an opinion survey, or within the detail of establishing a sampling plan of products in a manufacturing process. As another example, the population about which one wishes to draw inferences by sampling needs to be defined, in terms of details in the application domain, such as all voters in the state who will vote in the fall election or all voters in the county who are eligible to vote, or all products manufactured at a particular manufacturing site or product lines with high levels of consumer complaints.

Indeed, normative reasoning under uncertainty within a particular domain is dependent on the application details. For example, if one is going to conduct an opinion survey (or interpret the results of one) appropriately, part of the relevant knowledge is knowing factors within the domain that are relevant to ensuring a representative sample. If one does not apply appropriate domain knowledge in reasoning under uncertainty, then the knowledge of the math, out of the context of the application details, cannot deliver the normative reasoning leading to sound judgments for the situation.

In summary, domain knowledge and expertise are important contributors to normative reasoning under uncertainty within the domain. For example, important within medical science is knowledge and expertise in medical and pharmaceutical research, important within business is knowledge and expertise in market surveying and manufacturing, and important in opinion polling is knowledge and expertise in opinion research.

The close relationship between the math of probability and statistics, and domain knowledge in the areas to which it has been applied, is illustrated in the history of the development of the mathematics of probability and statistics. Gigerenzer et. al. (1989) have noted: "Perhaps more than any other part of mathematics, probability theory has had a relationship of intimacy bordering upon identity with its applications. Indeed, there was arguably no 'pure' theory of mathematical probability until 1933 …, and until the early nineteenth century, the failure of an application threatened the theory itself … For much of its history, probability theory *was* its applications" (pp. xiii-xiv). As probability theory spread to new domains, "…the mathematical tool shaped, but was also shaped by, its objects" (p. xiv). "From its beginnings in the mid-seventeenth century, probability theory spread in the eighteenth century from gambling problems to jurisprudence, data analysis, inductive inference, and insurance, and from there to sociology, to physics, to biology, and to psychology in the nineteenth, and on to agronomy, polling, medical testing, baseball, and innumerable other practical (and not so practical) matters in the twentieth" (p. xiii). The close relationship between the math of probability and statistics and the real world application domains to which it has been applied, is also evident in the cognitive realm: normative reasoning under uncertainty in real world situations incorporates and integrates both knowledge of the math of probability and statistics, and knowledge of the application domain.

Expert Consensus

As discussed in sections above, to apply the math of probability and statistics to a real world situation involving uncertainty, is not always the normative approach to reasoning for the situation. Sometimes the assumptions required to apply the math do not hold in the situation, so the math is not applicable; and sometimes, with either good or bad intentions, math knowledge is misapplied in a situation. Rather, what determines what is considered normative reasoning under uncertainty in various situations is the consensus of experts within the culture or relevant communities.

Among mathematicians in probability and statistics, there has not always been consensus. Gigerenzer et. al. (1989) have described the long and heated disagreement between giants of the field R. A. Fisher and the duo J. Neyman and E. Pearson regarding tests of significance and hypothesis testing, and note that the "vigorous controversies … have not ended" (p. 105). They added that "Disputes no less heated have characterized the relationship between Bayesians and frequentists" (p.105). Lopes (1982) reviewed literature related to the understanding of randomness, and raised the point that some assume "that randomness is clearly defined and well understood by those who are not naïve [mathematicians and philosophers]. Nothing could be farther from the truth" (p.628). She goes on to describe the different views of randomness of R. von Mises, G. Spencer-Brown, K. Popper, and others. Cohen (1981) made the point that the consensus of experts in a field can change over time, and cited the example of challenge arising to the Frege-Russell logic of quantification that seemed once "a universally received doctrine" (p.328).

As discussed in sections above, in reasoning about real world situations involving uncertainty, it is not just reasoning based on the math of probability and statistics that may apply, but necessarily also applicable is non-mathematical reasoning, such as heuristic reasoning, language interpretation, and the use of domain knowledge for the situation. Relevant experts in this case are both mathematical statisticians and experts in the domain of the application. Normative reasoning under uncertainty in the domain is determined by a consensus of these experts, if such a consensus is reached. If the math of probability and statistics is not clearly applicable within a domain of application, then normative reasoning under uncertainty in the domain using non-mathematical reasoning is determined by consensus of experts in the domain, again, if such a consensus is reached.

Normative Reasoning for Risk Management

Risk management involves reasoning under uncertainty and is an important function in business and industry. Industry standards for risk management have been established by industry-supported organizations, including the International Standards Organization (ISO) and, in the US, the federal government-sponsored Software Engineering Institute (SEI), both of which have assembled groups of experienced practitioners from industry (domain experts) to collaborate to establish standards for risk management, and to maintain those standards over the years. Established by domain experts and accepted by the community, the standards describe normative reasoning and behavior for risk management, and are used widely in industry as guidelines to promote effective practice of risk management.

ISO 31000:2009 Risk Management standards. The purpose of *ISO 31000:2009 Risk Management – Principles and guidelines*, as described on the ISO website, is to help "any organization regardless of its size, activity, or sector" to "increase likelihood of achieving objectives, improve the identification of opportunities and threats and effectively allocate and use resources for risk treatment." Within the standard, the risk management process includes risk assessment (identifying, analyzing, and evaluating risks) and risk treatment, all preceded by establishing the context for risk management. Risk analysis includes considering "the causes and sources of risks, their positive and negative consequences, and the likelihood that the consequences can occur" (p.21). A companion standard *ISO/IES 31010:2009 Risk Management – Risk assessment techniques*, "provides guidance on selection and application of systematic techniques for risk assessment" (Scope section, para. 1). To support identifying the risks that should be managed, the standard provides guidelines for selecting risk assessment techniques, and describes a set of 31 risk assessment techniques from which one may select, including brainstorming, structured or semi-structured interviews, Delphi Technique, check-lists, and Root Cause Analysis.

Regarding risk treatment, the ISO 31000:2009 standard provides a list of possible ways to manage or treat the risks one has identified:

- avoiding the risk by deciding not to start or continue with the activity that gives rise to the risk;
- taking or increasing risk in order to pursue an opportunity;
- removing the risk source;
- changing the likelihood;
- changing the consequences;
- sharing the risk with another party or parties (including contracts and risk financing); and
- retaining the risk by informed decision. (p.9)

The standard illustrates use of heuristic reasoning and the deep involvement of domain knowledge in the specification of the standard.

CMMI for Development, Risk Management standards. The SEI's Capability Maturity Model Integration (CMMI) for Development (Chrissis, Konrad, & Shrum, 2011) is a "collection of best practices that help organizations to improve their processes" (p.xv), originally established in its earliest form (CMM) for software development organizations in 1995, but also applicable to other organizations. CMMI-DEV includes a process area for Risk Management, with the purpose "to identify potential problems before they occur so that risk handling activities can be planned and invoked as needed across the life of the product or project to mitigate adverse impacts on achieving objectives" (p.481). The risk management practices include identifying and analyzing risks, including evaluating risk likelihood and risk consequence (i.e., impact and severity of risk occurrence) through human judgment, and tracking risks to monitor whether they have reached pre-planned thresholds to trigger management activities, including implementation of risk mitigation plans. The best practices illustrate the use of heuristic reasoning and involvement of domain knowledge, e.g., in the identification and quantification of risks and in the establishment of risk mitigation plans.

A Cognitive Framework for Normative Reasoning and Behavior under Uncertainty

In the previous sections, factors relevant to normative reasoning under uncertainty for real world situations, have been identified and discussed. In this section, the factors are integrated into a cognitive framework. The purpose of the framework is to present, in integrated form, key features of normative reasoning under uncertainty for real world situations, to further emphasize their role, and show how they tie together. The intended message is not that this is a complete list of features, but rather, a fundamental set of features that deserves emphasis, and provides useful perspective for researchers and educators. At the end of this section, the features and how they tie together are illustrated through an example.

Regarding the scope of the framework, it focuses on addressing human performance in real world situations involving uncertainty, in which what may happen or may be true, matters; and not with abstract, artificial, or simplistic tasks, e.g., as may occur with laboratory subjects being presented with short written problems to which to respond. In real world situations, the focus is on performing well or adaptively given the circumstances one is faced with and the consequences that may result. Real world situations are engaging, rich in detail, and call for action, and are where the value of normative reasoning under uncertainty delivers its payoff.

Real world situations may be categorized into "domains," e.g., involving medical science, business, politics, or consumer behavior. Often, within a domain, one can identify expert practitioners who may collaborate and reach consensus on what is normative reasoning under uncertainty for types of situations in the domain. Expert consensus provides the basis for identifying reasoning and behavior as normative. In contrast, to assume that any application of probability and statistics to a real world situation is normative, is unfounded. Also, to assume that every real world situation has an established normative standard for reasoning and behavior is unfounded.

For domains of application of probability and statistics, expert practitioners do exist, and provide a basis for establishing normative standards for reasoning under uncertainty within the domain, including normative application of probability and statistics. Still, there are controversies among experts, and so one cannot assume that there are normative standards for all applications of probability and statistics nor for all situations within a domain. What is and is not normative application of probability and statistics within a domain then has implications for what should and should not be taught, which is addressed in the last section of this paper on educational implications.

<table>
<tr><th colspan="5">A Cognitive Framework for Normative Reasoning and Behavior under Uncertainty</th></tr>
<tr><td>Steps:</td><td>Get initial situation understanding</td><td>Identify and evaluate applicable reasoning threads</td><td>Apply synthesizing reasoning and resolve understanding</td><td>Respond/ act, given resolved understanding</td></tr>
<tr><td>Actions (cognitive and behavioral):</td><td>• Get initial situation understanding including situation/ application detail
• Recognize elements/ sources of uncertainty
• Focus on question(s) about the situation</td><td>• Identify relevant information and reasoning threads to address the question(s)
• Seek out additional information about the situation to clarify or confirm the situation aspects related to reasoning threads; for mathematical reasoning, confirm that required assumptions hold</td><td>• Apply synthesizing reasoning to reasoning threads, leading to resolving understanding
• Seek out additional information about the situation to clarify or confirm the situation aspects related to synthesizing reasoning threads; for synthesizing using mathematical reasoning, confirm that required assumptions are satisfied</td><td>• Based on resolved understanding, respond/ act, e.g.: judge situation (e.g., make an estimate), choose an alternative, act to change the situation, continue to track situation to monitor uncertainties, abandon/avoid situation</td></tr>
<tr><td>Character-istics:</td><td colspan="4">• Normative reasoning is agreed upon by expert consensus within a domain of application and accepted by the community
• Reasoning incorporates and integrates: mathematical reasoning (based on logic, probability and statistics), non-mathematical reasoning (e.g., heuristic reasoning and language interpretation), knowledge of domain and situation/ application detail</td></tr>
<tr><td>Performance:</td><td colspan="4">• Proficient practitioners provide examples of normative reasoning and behavior
• For individuals not proficient in normative reasoning for a situation (reasoning based on expert consensus), in general, it is a personal norm to consider the risk of misapplying knowledge, in one's reasoning and behavior for the situation</td></tr>
</table>

Figure 1. A cognitive framework for normative reasoning and behavior under uncertainty

Having reviewed its intended scope, a cognitive framework for normative reasoning and behavior under uncertainty, expressing and integrating the factors discussed in previous sections, is presented in Figure 1 (above).

In the cognitive framework, the process for reasoning and behavior is broken down into four high level steps. The steps highlight the progression from initial situation understanding, to identifying and evaluating relevant reasoning threads, to applying synthesizing reasoning to the reasoning threads and resolving one's understanding, and responding/acting based on the resolved understanding. These steps may occur cyclically over time when a situation is being continuously monitored. This entire process is included in the scope of what is judged to be normative reasoning and behavior for the situation. For the entire process, mathematical and non-mathematical reasoning, and the use of knowledge of domain and situation/application detail apply. There are actions within the steps to seek out additional information to clarify or confirm one's understanding of the situation, including, for mathematical reasoning, to confirm that required assumptions hold for the situation.

Normative reasoning and behavior is presented in two forms within the framework. First, normative reasoning and behavior for a type of situation is agreed upon by expert consensus within a domain of application, and is accepted by the community. Proficient practitioners provide examples of this normative reasoning and behavior. Second, a personal norm for reasoning and behavior in the situation applies when an individual is not proficient in performing normative reasoning as settled by the expert consensus. For the person, there is a real risk of misapplying knowledge and reaping negative consequences, and so it is normative for the person to consider that risk in reasoning and behavior for the situation.

An example

Let us illustrate features of the framework, and of normative reasoning and behavior under uncertainty, with an example based on a real world situation. Consider the case of a hiring manager reviewing candidates for a job opening, with the purpose to determine which of the candidates to hire, or to hire none and keep looking. This situation involves uncertainty since the manager does not have complete knowledge of the candidates and how they may perform in the job. The situation involves risk, due not only to that uncertainty, but also because there are negative consequences to making a poor decision (such as hiring a candidate who ends up not performing well in the job, or who leaves the job after a short time, after company resources have been spent training him), and positive consequences to making a good decision (such as having a person exceptionally well suited to the job stay in the job, grow in the job, and serve the company well for years).

A first question is whether there is a normative standard for reasoning and behavior in this situation. Even though the situation of evaluating candidates for a job and making hiring decisions, is a common and important situation, there is no worldwide consensus on best practices for hiring people for jobs. However, there may be a consensus on such practices for a smaller community.

Let's say that Company A has developed a consensus on their hiring practices. For each candidate, they conduct a structured interview with the candidate and rate the interview; and they administer a test of job skills that are required for the job, and generate a skill performance profile for the person based on the test. To be hired, a candidate's job interview rating must fall within a specified acceptable range, and his skill profile must rate in a specified acceptable range for all required skills. If more than one candidate satisfies these conditions, the hiring manager judges, based on the interview and skill ratings, and his own judgment, the best candidate for the job. As defined, the hiring process involves heuristic reasoning. This is the hiring practice that has been in place for the past 5 years, and Company A may revise its practices in the future if improvements are identified and agreed upon. The company gets many of its job candidates from three educational/training institutions X, Y, and Z, and maintains historical records on the number and percentage of its hires that come from each of the three institutions. Over the past 5 years, new hires have been 25% from X, 50% from Y, and 20% from Z.

So, to follow the progression of the normative reasoning and behavior defined by Company A, using the steps highlighted in the cognitive framework, first, the hiring manager gets an initial situation understanding: he understands his typical goal to make a hiring decision, who the available candidates are this time, and the results of the structured interview and test of job skills for each of the candidates. Then the manager identifies and evaluates the relevant information to making the hiring decision (reasoning threads) for each candidate, considering each piece of information relevant to the decision. At this stage, the evaluation may lead to a candidate being eliminated from consideration because he does not meet the minimum requirements. Let's say that four candidates have acceptable ratings for the interview and the test of job skills. Now, synthesizing reasoning needs to be applied to consider all four remaining candidates and which is the best bet to hire. For this step, the company's normative process does not tightly constrain the decision, only saying that the manager is to consider the interview and skill ratings and also apply his own judgment. Consistent with these guidelines, the manager reviews in detail and compares the information for the four candidates. He eliminates two candidates who had mostly minimally acceptable ratings. Reviewing the full information on the two remaining candidates P and Q, the manager sees that one candidate Q has very high ratings for several skills, including a skill for which his other team members are not strong. He sees it as positive to be able to add someone strong in that skill to his team. In contrast, candidate P's ratings are all in the midrange. The manager decides to hire candidate Q.

A colleague, who knew that there were the two candidates P and Q in the final running for the job, suggested to the hiring manager that he hire candidate P, because P was from institution Y, which historically supplied 50% of new hires to the company, more than the 25% supplied by institution X, from where candidate Q had come. The hiring manager replied that there was no agreement in the company that that factor must be applied; and that he believed that one reason that institution Y provided more new hires was that more of their students were Spanish-speaking, which was a required skill for their jobs; and that candidate Q had that skill. The hiring manager also pointed out that, by basing hiring decisions on individuals' performance, and not on the institution where they were trained, the process was fair, not giving preference to candidates based on where they came from. The colleague listened with interest, and departed with a greater self-awareness of gaps in his own knowledge related to making a good choice for new hire.

The above example illustrates normative reasoning and behavior established by expert consensus, the use of heuristic reasoning and domain knowledge, and the richness of situation/application detail relevant to reasoning. In the example, there is mathematical information available on the base rates of new hires by educational/training institution, but that information is not required to be used in the process, and, in the example, reasons are provided for not using that information in the decision to hire an individual.

This hiring decision example structurally parallels examples that Tversky and Kahneman (1974) used to support their conclusion that people's use of heuristics, such as the representativeness heuristic, in reasoning under uncertainty, "lead[s] to systematic and predictable errors" (p.1131), such as ignoring base rates or prior probabilities of outcomes. In one of Tversky and Kahneman's examples, subjects were asked to judge the probability that a person has the occupation of engineer or lawyer, given a personality sketch as well as base rate information on the engineer vs. lawyer mix of the pool from which the sketch allegedly had been randomly selected. They found that subjects tended to use the base rate information if no personality sketch was provided; but tended to use only the personality sketch and ignore the base rate information when the personality sketch was provided to evaluate. The hiring decision example shows a situation where it is normative to ignore available base rate information. In a parallel fashion, subjects in Tversky and Kahneman's study ignore base rates in favor of focusing on the information provided about the particular individual being judged. Regarding synthesizing reasoning, Tversky and Kahneman (1974) presuppose that people should follow Bayes' rule, a straightforward number crunching, in making a judgment based on the combined evidence from the personality sketch and the base rate information. In contrast, in the hiring decision example, the synthesizing reasoning occurs in the domain of the real world, richly using situation detail, and is not just a numerical calculation in the realm of mathematics.

Implications for Educational Practice

The mathematics of probability and statistics provides a powerful tool for use in reasoning under uncertainty, in situations when it is applicable. There is also risk of misapplying the mathematics. A view of what is normative or appropriate reasoning under uncertainty, including normative application of probability and statistics, provides valuable guidance for the effective and adaptive reasoning and behavior in real world situations; as well as valuable guidance for educational practice regarding what should and what should not be taught. Such a framework for normative reasoning and behavior under uncertainty has been presented in this paper. Based on the framework, recommendations for educational practice in probability and statistics follow:

1. *Demonstrate the power of probability and statistics when used normatively/ appropriately, in different domains.* Provide real world examples that illustrate established normative standards for the application of probability and statistics in different domains, to show the power of probability and statistics when used appropriately. Use the established methods within the domain, which may integrate mathematical and non-mathematical reasoning, and provide the richness of detail involved in applying the methods.

2. *Use the domain of games of chance to teach probability and statistics.* Games of chance provided the original problems that mathematicians used in the first development of the math of probability and statistics (David, 1962; Gigerenzer et. al., 1989), and provide a well-established domain for the normative application of probability and statistics. Moreover, experience with the concrete instantiation of randomness in the common random phenomena associated with games of chance (e.g., the rolling of dice, and the blind drawing of balls from an urn), helps to build the mature understanding of random phenomena that is a foundation for understanding probability and statistics (Kuzmak, 2014).

3. *Don't teach students to ignore relevant domain knowledge and application/situation detail in situations involving uncertainty.* Appropriate use of probability and statistics is not just choosing math formulae that seem to fit a situation, and plugging in numbers, and calculating, regardless of application detail and whether required assumptions are satisfied. If one is teaching in a way that models that kind of behavior, then one is teaching students to reason inappropriately, specifically, to ignore domain knowledge and application detail that is relevant to appropriate reasoning. Statistics problem solving should regularly include analysis of the domain of application adequate to justify the applicability of the math to the situation, and to not merely assume its applicability.

4. *Explicitly teach the risks of misapplying probability and statistics.* To prevent students from misapplying probability and statistics, and from using information based on misapplication of others, explicitly teach common ways that probability and statistics can be misapplied, that they should avoid. Examples of such misapplication were provided in the above section on that topic, e.g., not ensuring a sample is representative of the population about which one wishes to make inferences, and not taking into account required assumptions when applying statistical techniques. Give examples of the negative consequences of misapplying probability and statistics, to emphasize the importance of avoiding such misapplications.

5. *Explicitly teach that the use of the math of probability and statistics is not always the best or normative approach to reasoning and behavior under uncertainty.* Give illustrations of normative reasoning under uncertainty in different domains, that use heuristic reasoning, and a combination of heuristic and mathematical reasoning. This provides further emphasis that the math of probability and statistics is a tool to be used when appropriate, which is not always. Give examples illustrating that there is not always an established normative approach to reasoning and behavior for a situation, and that there may be controversies over valid approaches, including controversies over how probability and statistics may apply.

Conclusion

The focus of this paper has been on providing a cognitive framework for viewing normative reasoning and behavior under uncertainty in real world situations, including situations involving risk. Such a framework provides a valuable reference for guiding what should be taught, and, in contrast, what should not be. Recommendations for educational practice in probability and statistics based on aspects of the normative framework have been provided in the section above. The framework for normative reasoning and behavior also provides a reference for evaluating whether particular examples of observed reasoning and behavior, identified as "errors" in the research literature, are rightly called so. Within the framework, a consensus of experts within the domain of application is the determinant of what is normative. The errancy of assuming that applying probability and statistics to a situation is normative, or assuming that every situation has an established normative response, has been noted.

The framework provides a fresh perspective on factors that are fundamental to normative reasoning and behavior under uncertainty, emphasizing the roles of mathematical and non-mathematical reasoning, and real world domain knowledge and situation/ application detail, and illustrating the roles of these factors with several examples rich in relevant detail. Within the framework, mathematics is viewed as a tool that may be applied appropriately or inappropriately, not a universal solution.

The risk of misapplying probability and statistics is identified as a significant risk at play in situations involving reasoning and behavior under uncertainty, and is illustrated through examples, discussion, and research findings. One faces this risk when one is not proficient at the normative reasoning and behavior agreed upon by domain experts. The framework recognizes that, in this case, it is a personal norm to consider the risk in one's reasoning and behavior in the situation.

References

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *The Behavioral and Brain Sciences, 4*, 317-370.

Chrissis, M., Konrad, M., & Shrum, S. (2011). *CMMI for development: Guidelines for process integration and product improvement* (3rd ed.) (pp.481-495). Boston: Pearson Education, Inc. http://cmmiinstitute.com/resources/cmmi-development-version-13

David, F. (1962). *Games, gods and gambling: The origins and history of probability and statistical ideas from the earliest times to the Newtonian era.* London, UK: Charles Griffin & Co. Ltd.

Garfield, J. & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*, 44-63.

Garfield, J. & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*, 592-596.

Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451-482.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). *The empire of chance*. Cambridge, UK: Cambridge University Press.

Gigerenzer, G. & Todd, P. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp.3-34). New York: Oxford University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, U.K.: Cambridge University Press.

International Standards Organization (ISO) (2009). *ISO 31000: 2009, Risk management – Principles and guidelines.* Geneva, Switzerland: International Organization for Standardization. http://www.iso.org/iso/home/standards/iso31000.htm

International Electrotechnical Commission (IEC) (2009). *IEC 31010: 2009, Risk management – Risk assessment techniques.* Geneva, Switzerland: IEC. http://www.iso.org/iso /catalogue_detail?csnumber=51073

Ioannidis, J. (2005). Why most published research findings are false. *PLoS Medicine, 2*, 696-701.

John, L., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*, 524-532.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus, and Giroux.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, U.K.: Cambridge University Press.

Kahneman, D. & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*, 582-591.

Kuzmak, S. (2014). What's missing in teaching probability and statistics: Building cognitive schema for understanding random phenomena. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. http://icots.info/9/proceedings/

Lopes, L. (1982). Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 626-636.

Lopes, L. & Oden, G. (1991). The rationality of intelligence. In E. Ellis & T. Maruszewski (Eds.), *Probability and rationality: Studies on L. Jonathan Cohen's philosophy of science* (pp. 199-223). Amsterdam: Rodopi.

Simmons, J., Nelson, L., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359-1366.

Tishkovskaya, S. & Lancaster, G. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education, 20*(2). Retrieved from www.amstat.org/publications/jse/

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.