5

Subjective Probability and Utility

Hank was feeling a little introspective. He, Betty, and Ralph had spent more time than any of them could afford trying to choose the best applicants from the pile. He had tried to behave as rationally as possible, but the information on the applications always left him in doubt—he never had a feeling that he was as well informed as he should be.

A while back he had told Betty that selecting new employees was a crap shoot. Apparently she thought he meant it was sheer chance. What he really meant was that there was a lot of uncertainty involved—he was convinced that uncertainty was a big part of the whole process for him.

On the other hand, uncertainty was not the only part of the process. It also was true that he tried to anticipate the outcomes that might result from hiring a particular applicant. It wasn't all that difficult. He merely tried to imagine the applicant in various situations that might arise and how he or she would handle them. Of course, he had very strong views about what was the right thing to do in those hypothetical problem situations. If it seemed like applicants would do what he preferred them to do, he liked them. If not, he didn't.

In decision research, riskiness, doubt, and uncertainty are captured by probability and value is captured by utility. Recall from Chapter 4 that two (EV and EU) of the four variations on the normative theory of choice use objective probability to represent the decision maker's uncertainty about whether the various outcomes actually will be attained if the option is chosen, and two (SEV and SEU) use subjective probability to represent this uncertainty. Similarly, two (EV and SEV) use value to represent outcomes' worth and two (EU and SEU) use utility.

❖ THE NATURE OF PROBABILITY

All four variations on the normative theory of choice assume that the probabilities have the mathematical properties required by probability theory—that is, that they are "real" probabilities and not just any old decimal numbers. This is equivalent to saying that uncertainty is appropriately represented as a probability.

While this assumption is not difficult to accept for objective probabilities, it seems rather shaky for subjective probabilities: Who knows where they come from? On the other hand, if subjective probabilities are not "real" probabilities, the SEV and SEU models lose their claims to normative (prescriptive) status. Because of this, a good deal of research in the 1960s through the early 1980s centered on whether those claims were valid by examining how well subjective probability conforms to probability theory. To understand this research we must briefly review the rudiments of probability theory.

A Brief Review of Probability

Probability theory is an abstract, axiomatic mathematical system of rules for assigning numbers to sets of hypothetical elements (Kolmogorov, 1950). As such, it has nothing to say about events in the real world—that comes later when the user of the theory ties its concepts to specific events of interest.

Mathematical probability theory begins with a set of hypothetical elements, consisting of individual elements (A, B, etc.), unions of elements ($A \cup B$), intersects of elements ($A \cap B$), and complements of elements ($A - B$).[1] A number can be assigned to each of these elements. The number assigned to an empty set of elements is .00 (which defines the lower limit of the range of acceptable numbers). The number assigned to a subset of elements is equal to the sum of the numbers assigned to each of its constituent elements (which defines additivity). The number assigned to the set of all elements is 1.00 (which defines the upper limit of the range of acceptable numbers). Thus the numbers assigned to the elements in question must lie between .00 and 1.00, and the system is additive.

Assignment of numbers to individual elements occurs when we attempt to use probability theory for real world applications. To do this we use one of three definitions.

The first definition is *necessary probability,* exemplified by the probability of a three being observed on a roll of a single die being equal to 1/6 = .17. That is, because a die (one of a pair of dice) has six sides, only one of which has three spots on it, and because we assume the die is unbiased and the roll is fair, the probability of .17 is necessary given the physical structure of the die. If we were speaking of a deck of cards, the probability of drawing a three would be 4/52 = .08, because there are four cards with threes on them in the deck of 52 cards.

A second definition is *frequentistic probability,* exemplified by actuarial tables used in the insurance industry. For example, by knowing the past relative frequency of thefts of the kind of car you want to insure, your insurance company can judge the probability that your car will be stolen—"The probability (past relative frequency) of theft of this kind of car is *X*." Then they can set the premium high enough to make it worth taking the risk of insuring you. Similarly, when a particular weather pattern is observed, the weather service can use the past relative frequency of rain under these conditions as the probability that it will rain this time—"The probability (past relative frequency) of rain under these conditions is *X*." Note the conceptual leap in going from the relative frequencies of past events to the probabilities of future events. It depends heavily on the assumption that however the process operated in the past, it will continue to operate in the same way in the future—thieves will continue to have the same taste in cars and climatic systems will continue to operate as they have before. Moreover, it assumes that the relative frequency for a collective of past events is applicable to a single future event (theft of your particular car or occurrence of rain today).

The third definition is *subjective (personal) probability,* exemplified by my statement that "I think the probability is about .75 that Senator Smog will run for President." This is a statement about my certainty (uncertainty) about future events. It is not at all clear where these probabilities come from, but they are quite distinct from necessary and frequentistic probability. There is no necessity for Senator Smog to run for President (except perhaps in his own mind), and it is not at all clear how relative frequency would apply. If he had never run before, the relative frequency would be undefined (because you would have to divide by zero), so does that mean that there is no way of assessing a probability that he will run in the future? If he ran in the last election, would it mean the probability of his running this time is 1.00? But even then, he only ran one time out of the something like 50 U.S. presidential elections, so does that make the probability that he will run again 1/50 = .02? Neither necessary nor frequentistic probability make any sense in this kind of situation.

Probability theory is deceptive in its simplicity. Exploring the logical implications of those few rules for assigning numbers to elements has permitted mathematicians to derive the useful tools of modern statistics. One aspect of these logical implications, *conditional probability,* has been of particular interest in behavioral research on probability judgment. To understand conditional probability, let us begin with an intuitive description, then move to a formal statement, and then explore its importance to research on subjective probability.

Conditionality and Bayes' Theorem

When the probability of one event is modified by the fact that some other event already has occurred, the modified probability is called a conditional probability. Thus Hank may think that the probability that a new job seeker will be successful is .50, if only because he does not know anything to sway his opinion one way or another. Then he reads the answer to the first question on the person's application. The information in that answer modifies Hanks uncertainty—making him less uncertain about success or more uncertain. The new probability of success is correspondingly higher or lower than the beginning probability—let us say that it goes up to .65. Then Hank reads the second answer and modifies the .65 either upward or downward, and so on until he has read the answers to all the questions on the application or until he becomes sufficiently certain about success to hire the person or sufficiently uncertain about success to turn the person down. That is, Hank's subjective probability (uncertainty) is conditional upon the answers in the application.

Formally, conditional probability is an algebraic consequence of the axioms we described above. Let us consider two events that, for reasons that will be clear in a moment, we will call *H* and *D* rather than *A* and *B*. The numbers (probabilities) assigned to elements *H* and *D* will be written as $P(H)$ and $P(D)$. Conditional probability, written $P(H \mid D)$, is defined in terms of the intersect of *H* and *D*, written $P(H \cap D)$ as well as $P(H)$ and $P(D)$. Thus the conditional probability of *H* given that *D* has occurred is defined as $P(H \mid D) = P(H \cap D) \div P(D)$, which is read as "the conditional probability of *H* given *D* is equal to the probability of the intersection of *H* and *D* divided by the probability of *D*." Conversely, the conditional probability of *D* given that *H* has occurred is defined as $P(D \mid H) = P(H \cap D) \div P(H)$, which is read as "the conditional probability of *D* given *H* is equal to the probability of the intersection of *H* and *D* divided by the probability of *H*." Multiplying both sides of the two equations by $P(D)$ or $P(H)$, respectively, and rearranging terms, yields:

$$P(H \cap D) = P(H \mid D)\,P(D), \text{ and}$$

$$P(H \cap D) = P(D \mid H)\,P(H)$$

Because the left sides of both of these equations are the same, the right sides must be equal to one another:

$$P(H \mid D)\,P(D) = P(D \mid H)\,P(H)$$

and, dividing both sides by $P(D)$,

$$P(H \mid D) = \frac{P(D \mid H)\,P(H)}{P(D)} \qquad \text{(eq. 3)}$$

Equation 3 is called Bayes' Theorem, after the Reverend Thomas Bayes (1958) who first recognized its primary implication. To make that implication clear, let us say that *H* stands for *hypothesis* and *D* stands for *data.* Starting on the right, we begin with the probability that a hypothesis is true before we procure data about it, $P(H)$, called the *prior probability.* Then we gather some data, *D*, and compute the probability that these data would have been obtained if the hypothesis were indeed true, $P(D \mid H)$, and divide it by the probability that these data would have been obtained whether or not this hypothesis were true. This fraction is called the *likelihood.* Then we multiply the prior probability by the likelihood to arrive at $P(H \mid D)$, the *posterior probability* that hypothesis H is true in light of the observed data, D.

Recalling Hank and the new job seeker, Hank's prior probability of .50 for the success was revised to .65 in light of the job seeker's answer to the first question on the application, indicating that the answer had a large impact on Hank's opinion—called the *diagnosticity* of the data, which is reflected in the likelihood. In effect, Hank had to ask himself, "How probable is it that this applicant would have given this particular answer if he were going to succeed, relative to the probability he would have given it whether or not he were going to succeed?"

❖ SUBJECTIVE PROBABILITY, BAYES' THEOREM, AND DECISION RESEARCH

Because the normative legitimacy of the SEV and SEU models rest on the legitimacy of subjective probability, for 20 years or more the agenda for many behavioral decision researchers was set by the question of how closely subjective probability conformed to probability theory. This work addressed three issues: measurement of subjective probability, subjective probabilities for simple events, and the revision of subjective probabilities in light of data.

Measurement Issues

Before you can do reasonable research on how closely subjective probability conforms to probability theory, you have to decide how to measure subjective probability. Clearly, if different measurement methods yield different results, it is going to be difficult to evaluate conformity. This is not the place for a tutorial on methodology; suffice it to say that many methods have been tried (direct assessment, psychophysical measurement, inferences from bets, confidence ratings, and verbal statements such as "sure thing" or "toss up"). Early on, comparisons were made among these (e.g., Beach, 1974; Beach & Phillips, 1967; Beach & Wise, 1969; Galanter, 1962; Wise, 1970). Of course, the measurement technique one uses is going to be dictated in part by the demands of the research setting and task, but based on the research, and judging from the frequency with which researchers use it, direct assessment appears about as good as any other method (von Winterfeldt & Edwards, 1986). That is, simply asking people to give a number to represent their opinion about the probability of an event appears to produce data that are not markedly different from data produced by the other methods, and the procedure is simpler than for most other methods. The least attractive method involves verbal statements—individuals may be consistent in what they mean when they use statements like "sure thing," "very likely," "unlikely," but there is very little agreement across individuals about the level of probability represented by the statements (Lichtenstein & Newman, 1967). The irony of this is that many experiments, both in behavioral decision research and in other areas, rely on rating scales anchored with verbal statements of this kind.

Direct assessment can be done in two ways. One way is to ask people to state the probability. The other way is to ask them to state the relative frequency (or proportion) with which an event might be expected to occur ("The probability of a given person having characteristic X is .10" versus "Ten people out of 100 can be expected to have characteristic X" or "Ten percent of these people have characteristic X"). For a long time it was assumed that all direct assessments were equivalent, but as we will see later, this may not be true, and what one concludes about the nature of subjective probability depends upon which kind of assessment, stated probability, or relative frequency (proportion) one asks decision makers for.

Accuracy and Coherence for Simple Events

One way of evaluating the conformity of subjective probability to probability theory is to study areas in which it is possible to calculate objective probabilities (necessary or frequentistic) and then compare decision makers' assessments with them. This is called *accuracy.* It frequently is found that people tend to give assessments that are a bit too high for objectively low probability events and a bit too low for objectively high probabilities (e.g., Preston & Baratta, 1948). This generalization ignores many exceptions, but as a summary it is roughly correct.

Sometimes participants' assessments bear almost no discernible relationship to objective probabilities, but that does not mean that they do not conform to probability theory. It is quite possible for a person's probabilities to be *coherent* (i.e., interrelated in the ways demanded by probability theory) even if they are inaccurate—it merely means that the person is not well informed about the necessary probabilities or the relative frequencies for the domain in question. Ignorance is not the same as incoherence.

Coherence among subjective probabilities that are inaccurate, or for which there are no objective counterparts, can be measured by having participants assess the probabilities for each of the events and each of the compounds (unions, intercepts, or conditionals) in a set. Then the experimenter analyzes the degree to which the assessments "fit together" in ways dictated by probability theory.

For example, Peterson, Ulehla, Miller, Bourne, and Stilson (1965) presented participants with a list of personality traits and for each asked a question of the form, "How many people in a hundred are witty, brave, and so on," for $P(A)$, $P(B)$, and so on. Then they asked for conditional probabilities with questions of the form, "One hundred persons are known to be brave, how many would you expect to be witty?" for $P(A \mid B)$ and, "One hundred persons are known to be witty, how many would you expect to be brave?" for $P(B \mid A)$. Recall from our previous definition of conditionality that $P(A \mid B)\ P(B) = P(B \mid A)\ P(A)$. Therefore the product of the participants' assessments of $P(A \mid B)$ and their assessments of $P(B)$ ought to be equal to the product of their assessments of $P(B \mid A)$ and their assessments of $P(A)$. To test this, Peterson and his colleagues merely correlated the two products across participants; the mean correlation was .67. This may not seem very high, but the experimenters recognized that asking people to do a strange task like this is unlikely to produce very stable assessments, so they measured the reliability of the assessments; the mean reliability correlation was .72. Because response unreliability (.72) places an upper limit on coherence, the correlation of .67 for the latter is quite encouraging. Even higher coherence for assessments of familiar, concrete events was obtained in other studies. For example, Barclay and Beach (1972) obtained correlations in the high .70s, .80s, and .90s both for group data and for individual participants.

Accuracy and Coherence in Probability Revision

Recall that Bayes' Theorem follows directly from the definition of conditional probability, and it can be interpreted as a mechanism for revising probabilities in light of data. As a result, it affords a way of examining the accuracy and coherence of subjective probabilities in a more dynamic situation than for single events. Edwards and his colleagues were the first to do "Bayesian" studies of accuracy and coherence. They used variations on what was called "the bookbag-and-poker chips task." This consisted of showing participants two or more cloth bags (bookbags were the forerunners of backpacks as the conveyance of choice for students' textbooks). Each bag contained a mixture of blue and red poker chips. The proportion of blue chips differed from one bag to another, and participants were told the proportions for each bag.

Out of the participants' view, the experimenter randomly selected one of the bags, drew a sample of chips from it and told participants the proportion of blue chips in the sample. Then, for each of the bags of chips, the participants assessed the probability that that bag was the one that had been selected.

Because the bag to be sampled was randomly selected, the number of bags determined the prior probability for each being the selected bag: If there were two bags the prior probability for each was .50, if there were three the prior probability for each was .33, and so on. The proportion of blue chips in the bag determined the likelihood; if a bag's proportion was high and the sample had lots of blue chips in it, the likelihood was high that the bag was the one that had been selected—if the proportion was low but the sample had lots of blue chips in it, the likelihood was low that the bag had been selected. Thus the posterior probability (each bag's probability of having been selected in light of the composition of the sample of chips) was jointly determined by the prior probability and the likelihood, and this should have been reflected in the probability assessments given by the participants for each of the bags.

There were two major findings (Phillips & Edwards, 1966). First, participants tended to treat the prior probabilities of the bags as if they were more equal than they actually were. Second, participants tended to be less influenced by the data (the blue chips in the sample) than they should have been. As a result, their posterior probability assessments were *conservative* relative to the posterior probabilities that a statistician would arrive at using Bayes' Theorem. Close examination of the data showed that participants' posterior probabilities after the first data were moderately accurate, but that successive draws from the same bag (remember that the posterior after one observation of data becomes the prior for the next observation of data) led to increasingly severe conservatism (Peterson, Schneider, & Miller, 1965).

Conservatism was robust in that it was obtained in many replications of and variations upon the bookbag-and-poker chips experiments, and in even more realistic tasks of comparable logical structure. Although training sometimes reduced conservatism in a specific task (e.g., Christensen-Szalanski & Beach, 1982; Peterson, DuCharme, & Edwards, 1968; Wheeler & Beach, 1968), the general conclusion remains that, compared to Bayes' Theorem, decision makers' revised subjective probabilities are neither accurate nor coherent.

❖ REEXAMINATION OF SUBJECTIVE PROBABILITY

By the late 1980s, it was generally agreed that probability theory does not adequately describe subjective probability—or, put another way, subjective probability does not conform to probability theory. However, most investigators persisted in their belief that the four variants of the expected value model were descriptive of choice, and subjective probability is a component of two of those variants. One might think, therefore, that these researchers would move from testing the fit with probability theory to investigating the nature of subjective probability itself, in the hope of salvaging the SEV and SEU models. This is not what happened. For the most part, interest in subjective probability simply evaporated. It disappeared from the agendas of conferences, both in the United States and abroad. It was as though everyone simply was bored with the topic and anxious to move on to something else. And then a new voice, with new data, was heard calling for a reexamination of our conclusions about subjective probability judgments.

The new voice belonged to Gerd Gigerenzer. In the late 1980s he and his colleagues undertook a reexamination of the literature on subjective probability. Their conclusion was that it is necessary to differentiate between decision makers' assessments of relative frequency, which implies long-run probabilities of events, and assessments of the probabilities of unique events (Gigerenzer, 1991).

It has long been known that people are very good at assessing proportions (Peterson & Beach, 1967), and Gigerenzer's studies show that when decision makers make assessments in the form of relative frequencies or proportions, many of the problems with subjective probability are greatly reduced or disappear. However, when asked for probabilities for unique events, the problems are strongly in evidence. Gigerenzer's conclusion is that frequentistic judgments made by people who are reasonably familiar with the domain of interest are apt to conform reasonably well to the demands of probability theory. Probability assessments for unique events are not very apt to conform to probability theory. This conclusion makes sense in light of many statisticians' strong opinions that applied probability theory only addresses long-run relative frequencies, and has no meaning for single events (e.g., von Mises, 1957).

In fact, it has been clear for quite a long time that it is necessary to differentiate between judgments for long-range and for unique events (Lopes, 1981). Beach, Barnes, and Christensen-Szalanski (1986) proposed that decision makers use different *judgment strategies* for different *judgment tasks* encountered in different *judgment environments,* and that the final judgment is *contingent* upon all three. The strategies fall into two categories, aleatory and epistemic. (An aleator is a dice player, hence aleatory refers to necessary and frequentistic probability. Epistemology means knowledge, hence epistemic refers to the use of knowledge to derive subjective probabilities.) The general idea is that the decision maker selects one or the other strategy depending upon how he or she frames the judgment task (whether chance appears to be an important component, whether repeated versus unique events are involved, whether statistical or causal logic is the norm for the domain in question). He or she then applies the selected strategy with more or less rigor depending upon the demands of the judgment environment (the payoff for accuracy, whether the judgment can be revised later, the degree to which the decision maker's credibility is on the line, the quality of the information with which he or she must work).

More recently, Gigerenzer, Hoffrage, and Kleinbolting (1991) proposed a process model for both aleatory and epistemic judgment based on work by Egon Brunswik. Thus things have come full circle; we began with Brunswik in Chapter 3 and we seem to have come back to Brunswik.

❖ THE NATURE OF UTILITY

We now turn to the fit between subjective worth and utility theory. Unfortunately, the research on subjective worth does not provide as clear a test of its fit with theory as the research on subjective probability provides; the evidence is more indirect.

As we have seen, the methods for comparing decisions with normative prescriptions are inherent in the nature of the normative theory itself. Because prescriptive theory views choices as gambles, and the usual prescription is to maximize expectation, it is widely accepted that studying choice means studying how people deal with gambles. As in Bernoulli's example in Chapter 4, the decision about sending a ship to the New World, this can consist of deciding to take or not to take a gamble (send the ship or not), or deciding on a fair price for a gamble (insurance). Or, as in the example of the lumber executive and Mount Saint Helens, it can consist of choosing the best from among two (or more) bets. In the laboratory, participants might be presented with a pair of gambles for which the expected value of gamble M is greater than that of gamble N, in which case they are predicted to choose M. In successive presentations, the payoffs can be systematically changed so that at some point the expected value for gamble M becomes less than that of gamble N. Of interest is the point at which decision makers switch from preferring gamble M to preferring gamble N, from which, with a little algebra, one can infer the utilities underlying their choice of one bet over the other.

❖ UTILITY THEORY

Just as formal probability theory is a way of assigning numbers to events, and not a theory about decision makers' uncertainty, utility theory is another way of assigning numbers to events and not a theory about what is valuable to decision makers. In application, however, utility theory is used to represent preferences among potential (or obtained) outcomes of a decision, and the question is how usefully it does its job.

As Yates (1990) has pointed out, there are two ways of relating preference to the "objective" value of outcomes. The first is called a *value function,* which represents the increase in the strength of the decision maker's preferences as a function of the outcomes' objective value. It is as if there were a scale in the decision maker's head on which the various outcomes are placed, such that the ordering of their locations are consistent (higher scale values mean higher preference), and the distances between the ordered outcomes on the scale represent meaningful differences in preference for the outcomes (the scale is ordinal). This first kind of scale is the most common view of utility—the relative preference of various outcomes.

The second way of relating preference to the objective value of outcomes is called a *utility function.* Here the assumption is that preference reflects both the value of the outcome to the decision maker *and* his or her feelings about risk (i.e., uncertainty about whether the outcome will or will not occur). Using conventional expected value logic, this means that preference is for gambles rather than merely for the outcomes. Hence it is as if there were a scale in the decision maker's head on which the various gambles are placed such that the ordering of their locations is consistent and the distances between them are meaningful. This second kind of scale is the one used in most discussions of utility in decision theory and research.

There are numerous versions of utility theory (von Winterfeldt & Edwards, 1986), but they all make three fundamental assumptions:

> *Connectivity.* They assume that the decision maker can judge his or her preferences (or indifference) when faced with two gambles.
>
> *Transitivity.* They assume that preferences among gambles are consistent such that if gamble M is preferred to gamble N, and gamble N is preferred to gamble O, gamble M is preferred to gamble O.
>
> *Summation.* They assume that the preference for a gamble is greater than the preferences for any of its component parts. For example, the preference for a gamble that offers a payoff of $50 and a movie ticket must be greater than the preference for the $50 alone or for the movie ticket alone. That is, the preference for a compound outcome of a gamble is a combination (usually the sum) of the preferences for the component outcomes.

If these and some ancillary enabling assumptions are met, it can be formally shown that gambles can be arrayed according to preference on an underlying scale of utility.

Because utility theory is an abstract method of attaching numbers to events, it is not altogether meaningful to talk of testing it. However, it is meaningful to talk of testing the degree to which decision makers' preferences conform to its assumptions and implications, and it is here that the behavioral decision research has focused, but—to repeat—the tests, which are far fewer than one might expect, have been less direct than in similar research on probability theory.

Research Results

- If the connectivity assumption holds, decision makers' preferences ought to be robust because connectivity means that they know what they prefer. Instead, it is found that preferences change depending upon task characteristics, measurement methods, context, and the probabilities with which they are associated (e.g., Fischhoff, Slovic, & Lichtenstein, 1980; Fryback, Goodman, & Edwards, 1973; Schoemaker, 1980; Slovic & Lichtenstein, 1983).

- If the transitivity assumption holds, the order of decision makers' preferences should be reliable. Research shows that this frequently is not the case and that intransitivity is easily induced (e.g., Tversky, 1969).

- If the summation assumption holds, preferences for compound gambles ought to be a function of the sum of their component gambles. Again, research finds that this is not always the case (e.g., Shanteau & Anderson, 1969).

To convey the flavor of the research that leads to these conclusions, consider a study by Tversky (1967). Using inmates in a federal prison as participants, Tversky asked each inmate to state the price he would ask to sell his right to play a particular gamble. Everyone in the room was given the same gamble, and the idea was that each inmate should ask a price lower than anyone else's so they, and not someone else, could sell the gamble to the experimenter. (Because the experiment was set up so that the inmates might not get the opportunity to play their gamble, in which case it became worthless, it was better to sell it and make at least the sale price.) On the other hand, they should not sell the gamble for less than the worth of the gamble (i.e., for less than its expected value).

The gamble's probabilities were presented as a pie diagram with a spinner attached to the center. If the spinner landed on one section of the pie, the inmate who got to play would win some stated amount, and if it landed on the other section he would win nothing. Thus the expected value of the bet was the product of the probability of winning, from the pie diagram, and the inmate's value for the payoff. The payoffs were in cigarettes and candy, which were used by inmates as currency within the prison. There were simple gambles and compound gambles—the latter had compounds of the payoffs that had been offered for some of the simple gambles and were designed to test the summation assumption described above.

The presumption was that the competition to sell the bet would drive down the inmates' selling prices until the lowest price asked would be equal to the subjective expected utility of the gamble. That is, $\$_s = (P \times V)$, where $\$_s$ is the lowest price any inmate in the group asked for the bet (in cigarettes or candy), P is the probability of winning (from the pie diagram), and V is the inmate's private value for winning whatever has been offered as the payoff. Because Tversky knew the probability of winning and the price asked by the inmate who under-priced everyone else, he could figure out what that inmate's value must have been for the payoff:

$$\$_s = \mathrm{P} \times \mathrm{V},$$

$$\mathrm{V} = \$_s/\mathrm{P},$$

which is to say, the price divided by the probability reveals the inmate's value for the payoff.

To state the complex results rather simply, Tversky found that inmates' asking prices suggested that they in fact evaluated the gambles in terms of the product of the probability and the value of the payoff, and that value appeared to be additive, both of which are congruent with utility theory. However, the inmates consistently set selling prices higher than the expected value of the gambles. In any strict sense, utility theory does not allow for what appears to be the inmates' desire for a profit margin or, interpreted another way, their value for retaining the gamble (a value for gambling). On the other hand, common sense finds both of these explanations reasonable.

Utility theory generally assumes that the value for a payoff must be the same whether it is a "sure thing" or whether it is part of a gamble. To test this, in one condition of the experiment Tversky's inmates simply set a price on each of the payoffs. Later, when these same payoffs were then included in gambles, the inferred value for them (using the equation) was not the same as the simple prices. This result contradicts the utility theory assumption and implies that value is not independent of risk, which means that the simple expected value equation is not an adequate description of the determinants of participants' utility for the gamble.

This is but one of a number of studies that obtain data suggesting that utility theory is not a very good description of human preferences, even in well-controlled experimental conditions.

❖ SUMMARY

We have been examining the two components of the expected utility model: subjective probability and utility. We began with the mathematical formulation of probability and then moved on to examine Bayes' Theorem, which is a consequence of the definition of conditionality. Behavioral studies comparing subjective probability with the demands of formal probability theory show that some similarity often is found for simple events but that there is consistent error for more complex events and for the revision of subjective probability.

After a hiatus, research on subjective probability returned in a new line of work by Gigerenzer (1991) that claims to refute many of the negative conclusions previously reached. It appears that both accuracy and consistency may be linked to decision makers' use of different strategies for assessing subjective probabilities, depending on the demands of the problem and the environment in which the problem is encountered.

Finally, we examined some of the basic assumptions of utility theory and described the general results of attempts to evaluate its adequacy as a description of human preferences. In fact, it does not come off too well. All in all, what with the unsettled question about whether probability theory adequately describes human uncertainty, the failure of utility theory makes acceptance of the expected value approach to choice highly tenuous.

❖ NOTE

1. Recall that the union of elements is both *A* and *B*, that the intersect is either *A* or *B*, and the complement is what remains when *B* is subtracted from *A* (or vice versa).