**MAKING GOOD CHOICES: AN INTRODUCTION TO PRACTICAL REASONING**

## CHAPTER 5: RISK AND PROBABILITY

Many of our decisions are not under conditions of certainty but involve risk. **Decision under risk** means that the agent is not practically certain, for each option, that the intended outcome will happen, but can estimate the chance that it does. The agent, in a risky choice situation, realizes that there is some chance that one or more alternative states might be in place, or that the required state could change before the intended outcome happens. If either should be the case then the intended outcome will not happen but another one will, perhaps one that threatens the agent's goal. When risk is real, the agent *knows* that any one action has two or more possible outcomes, depending on which state will exist when the action is done, but does *not know* which one it will be. Thus, the outcome is uncertain.

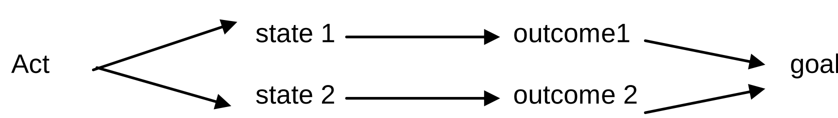
The simple example of flipping a coin provides a convenient and clear model of choice under risk. For contrast we will make the agent's menu of options mixed, containing a certain and a risky option. Let's suppose an agent has a goal of making money and is offered a choice between (a) $10 for sure or (b) the flip of a fair coin with $25 as outcome for heads and $0 as outcome for tails. The agent is certain (assuming the agent can trust the offer) that option (a) has a $10 outcome, but is uncertain about the outcome of option (b); it is uncertain to the agent that heads will land up and the agent knows that there is an alternative – tails – that threatens the goal. But at least the agent can estimate the degree of uncertainty that $25 will be gained in option (b).

Risk is the agent's uncertainty about the outcome of an option, when the degree of uncertainty is known or can be estimated. Uncertainty is a very important factor in practical reasoning, and will be represented by an estimation of the probability that the required state exists versus the probability that an alternative state exists. By formally representing risk as probabilities, several things are accomplished. First, it makes it difficult for an agent to neglect or forget the risk factor when it is a decision under risk. Second, it makes the agent think about the degree to which risk might detract from the utility of the outcome of an option. Third, for an agent who might be unduly afraid of risk, or one who might be thrilled by risk, it provides a moderating counterbalance toward realistic risk estimates.

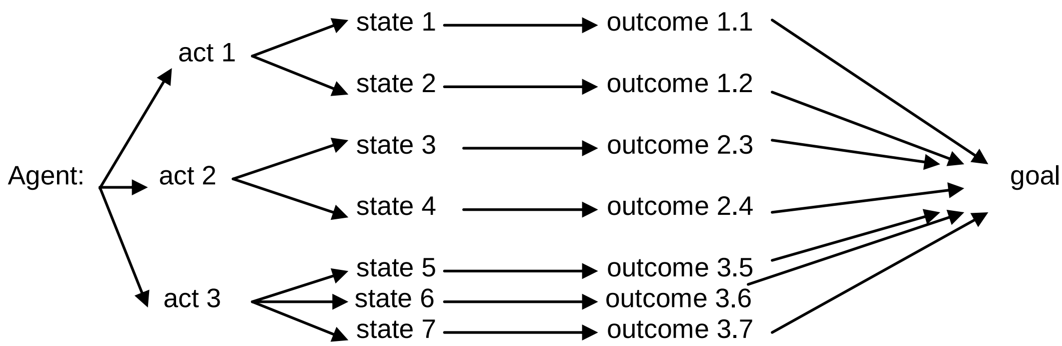
In this chapter we will set up the method by which risk will be represented in decisions. Then, in the next chapter we will apply this method to various risky decision problems.

5.1 Risky decisions

Where is risk located in a decision? Risk is a feature of the belief an agent has about the required state-of-the-world, it is not a feature of the desire the agent has for the goal. In risky decisions, the agent is less than maximally certain that the state is (or will be) in place; the agent is uncertain to some degree about the state, but can estimate the likelihood or probability that the required state exists. This means that in risky decisions the degree of uncertainty is greater than zero (zero or no uncertainty would mean that it is a case of certainty) but less than total (total uncertainty would mean complete ignorance). In risky decisions, the agent belkieves that there is a real possibility that some alternative state, one or more, might exist, in which case an act the agent might decide to do would have an alternative outcome. These other possible outcomes could affect the goal negatively, distancing it from the agent and perhaps even destroying any opportunity to achieve the goal at all. The basic structure for a risky option is:

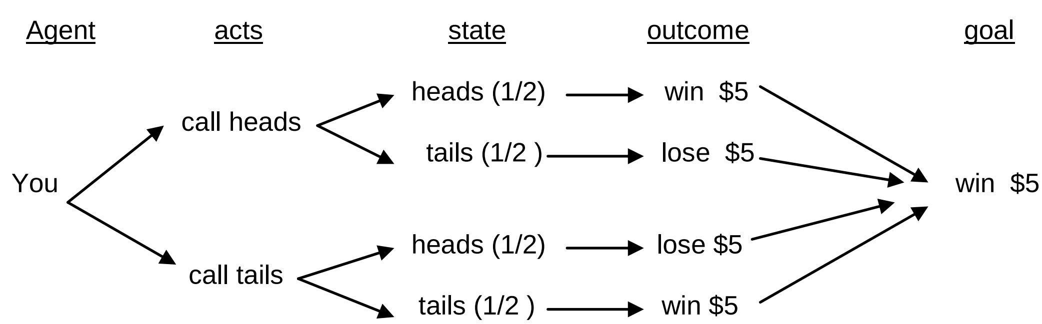



Because one of these outcomes might threaten the agent’s goal, **outcome, for decisions under risk,** is defined as the sub-set of all consequences of an option that changes the level of goal achievement or degree of goal fulfillment, positively or negatively. If we expand the basic structure for decisions under risk to, say, three options, we have this diagram:




In risky decisions, the idea is that the agent examines the alternative states for each option and tries to estimate the likelihood that one or another state will occur. On the basis of the likelihood, the agent forms a reasonable degree of expectation that doing an act will produce the intended outcome.

Flipping a coin provides a convenient and familiar way to illustrate decisions under risk. Suppose someone offered you the following little gamble. They flip a fair coin, and you call heads or tails. If you call it correctly, you win $5. If you call it incorrectly, you pay the person $5. Let's say that your goal is to win $5. We can easily fill in the above decision structure to get this:




There are several things to note. First, each time you accept this gamble the best you can expect as an outcome is $5 and this equals your goal. The best outcome is the upper limit of what you can rationally hope for in a risky decision; your **hope limit**. Second, the worse you can expect is to lose $5. This is not only the complete loss of your goal, it is also a payment you must give. This is your **security limit** in a risky decision; you have no reason to fear that any worse outcome can happen in the decision situation. Third, you can form a reasonable degree of expectation of achieving your goal for each option. In this case, you would be reasonable to expect a $5-gain outcome (your hope limit) no stronger and no weaker than the strength of your expectation of a $5-loss outcome (your security limit). Equal expectations in this case are based on the fact that for each option the state that promotes and the state that frustrates your goal achievement can occur with equal likelihood (1/2). Most risky decisions, as we will see, require unequal expectations. With this example in mind, we now want to make the idea of an agent's degree of confidence or expectation, both under certainty and under risk, more clear than it has been thus far presented.

## 5.2 Belief and belief strength

Both certainty and risk are connected to the agent's beliefs, not the agent's desires. We will accept a theory of belief that holds that beliefs can have a range of possible strengths. Someone might hold a belief very strongly, another person might have the same belief but hold it weakly, and a third person might believe the same thing with a strength falling somewhere between strong and weak. Take, for example, the statement: There is extra-terrestrial life in our universe. There are those who believe this very strongly, and others who disbelieve it equally strongly. Most people, however, are not so sure of its truth and have some doubt. They are uncertain, to a degree, of its truth. It is not that they disbelieve the statement, which would mean that they take it to be false. And they are not neutral, refusing to believe it to any degree, as well as refusing to disbelieve it to any degree. Rather they believe it, but not with the maximum degree of strength. It is quite common for a person to state what they believe and then add the phrase "but that's just my opinion" to indicate that they are not completely confident that their belief is correct. Strong believers will not add such a phrase to their belief; strong believers will demand a lot of negative evidence before they give up a belief, while those whose beliefs are less strong will more easily give it up.

We now stipulate that in every case of belief there is a **degree of confidence**, and the degree of confidence is the strength with which a believer holds a belief. At one extreme there is the maximum degree of confidence possible; this is certainty. At the other extreme there is the maximum degree of doubt possible; this is disbelief. In between these two extremes there are many different degrees of confidence with which a believer might hold a belief. A common way we have to get a person to reveal his/her degree of confidence in a belief is asking how much the person is willing to bet that her belief is true. If she answers, "I'll bet my life on it.", then we know her degree of confidence is maximal: certainty. If she answers, "I won't bet anything at all, all bets are off.", then the degree of confidence is minimal. If she answers, "I'm willing to bet $10 (my car, my house, all my savings…) that I'm right.", then we know she has a degree of confidence between maximal and minimal and a rough estimate of its strength is the value of what she is willing to bet.

Given this definition of degree of confidence, we must now ask: for any given belief, what would be a *reasonable* degree of confidence for a believer to have? What determines the degree of confidence that a person *should* have, and a rational person would have, with respect to a belief? In the theory of belief we are accepting, it is the evidence the person has that their belief is correct that determines what level of confidence is reasonable. As a general rule: <u>the more evidence a person has that a belief is correct, the stronger the degree confident the person can hold that belief.</u> Likewise, the poorer the evidence, the less confident believer should be about the correctness of their belief. Suppose, for example, that you are told that you lost $5 because you called tails and the coin landed heads. However, you did not see the coin land. You ask on what basis you should believe that the coin landed heads. The person tells you that he heard it from a friend, who heard it from another friend, who was told by the person who flipped the coin. Is this good evidence? With what strength would you believe that the coin landed heads? How ready would you be to hand over $5? You would probably have doubts, and quite reasonably not be very confident in believing the statement. But now compare this to the evidence you would have if you directly observed the coin land heads. In this case your degree of confidence that the coin landed heads would no doubt approach certainty. A **reasonable degree of confidence**, then**,** is one that matches, that is proportional to, the amount of evidence the believer has that the belief is correct.

5.3 Belief and risk

Given these concepts of belief and reasonable degree of confidence, it is now time to return to decisions under risk to see how the element of risk is represented within the decision framework. Using the decision diagram above containing 3 options, 7 states, and 7 outcomes (section 5.2) you will see that for each option there are two or more alternative states-of-the-world, each leading to different outcomes. Each states has a certain chance of happening; we will say each has a probability of existing. The agent, however, though certain that any given option from the menu will yield *an* outcome, is uncertain *which* one it will be. The agent is uncertain which possible outcome an option will yield because the agent is uncertain which possible state-of-the-world is or will be in place if that option is chosen.

Even though uncertain, in decisions under risk (by definition) the agent can form reasonable degrees of confidence about which outcome. And this means that the agent has some evidence concerning the alternative possible states-of-the-world; otherwise reasonable degrees of confidence could not be formed. Because the agent cannot be certain about the state (for example, no one can be certain that a fair coin flipped will land heads) the evidence must be the probability that a given state exists. So: in decisions under risk, the agent's degree of confidence that an outcome will result should reflect the evidence that the required state exists or will exist, and this means that an agent's degree of confidence that an option will yield a given outcome should be no more and no less than the probability of the required state. To return to our convenient little model of flipping a coin for $5, you should not expect a $5 outcome from calling heads with more confidence than the probability of heads landing up, both of which is ½. The general rule is:

- The more probable a state (S) is, the more confidently the agent should believe that the option containing S will yield the outcome requiring S.
- The less probable S is, the less confidently the agent should believe that the option containing S will yield the outcome requiring S.

An agent's degree of confidence about the outcome of an option is rational, then, if it accurately and reliably reflects the probability with which the state required for that outcome happens. In framing a decision under risk, a number will be used to represents the agent's degree of confidence, and that number should match the probability that the required state exists.

### 5.4 Probability

How is evidence concerning probability acquired? There are two basic sources. First, there is the actual experience and the records of the rate of frequency of a given state happening. This is a matter of statistics and gathering real data. We will group this under the heading: factual probability. Second, probabilities can sometimes be calculated by counting possibilities out of a larger set of possibilities. This comes under the heading: pure probability. These two sources will typically allow an agent to estimate and assign initial probabilities to states. Once initial probabilities are assigned to states, some basic rules of the probability calculus will allow the agent to combine initial probabilities. An agent's degree of confidence about the outcome of an option should accurately and reliably reflects either the factual or the pure probability with which the state required for that outcome happens. Before turning to the method of assigning initial factual and pure probabilities, here are some examples to help make these ideas clear

Example 1: Suppose that you have evidence from the past records that 6 out of every 10 automobile traffic accidents involve no injuries to anyone, only vehicle damage. What is the rational degree of confidence with which you should believe that no one will be hurt if you get into an automobile traffic accident? Answer: .6 degree of confidence is rational, based upon this (hypothetical) statistical record.

Example 2: Someone rolls a fair die. You stand to win $50 if the side with two dots lands up. With what degree of confidence should you believe: I am going to win $50? Here is the evidence you need to form a reasonable degree of confidence. There are six possible sides to a die, each with a number of dots from 1 to 6. Each side has an equal probability of landing up, for the die is fair. The desired side with two dots is, thus, one possibility out of six possibilities. 1/6 = .17. So, your degree of confidence should only be .17 you are going to win $50

Example 3: What is the rational degree of confidence that a fair coin flipped will land heads-up? The coin can land in only two possible ways, heads being one of the ways. Thus, a reasonable degree of confidence is .5.

5.4.1 Estimating initial probabilities.

We will first look at the category of **factual probability**, also called "statistical probability" or "empirical probability". This is probability based on past frequencies of events, about which observations have been recorded, statistics kept, or a reliable memory (history) exists. Our assumption is that the agent has access one way or another to the record of past frequencies. Let's take a typical example. Suppose someone asked the question: what is the probability of a person who smoked one pack of cigarettes a day for 10 years getting chronic respiratory illness? There are two things to identify in this question. First, there is a specific **property or event of interest**. In this case it is: getting a chronic respiratory illness. Second, there is a set of properties or events within which the property or event of interest is to be located. In this case it is: all people who smoked one pack of cigarettes a day for 10 years. This large set of events within which the event of interest is located is called the **sample space**. So, our question can be rephrased to this: Of all the people who smoked one pack of cigarettes per day for 10 years (count them up and this = the sample space), how many have had chronic respiratory illness (from the sample space count these up and this = the event of interest)? The records have to be consulted. Perhaps government agencies have kept statistics, perhaps the medical profession is doing a long-term study on just this issue, or perhaps the tobacco industry is amassing observations in anticipation of future legal battles. Whatever the source, we will end up with two numbers. Suppose the records reveal that, say, 150,000 out of 600,000 (=1 in 4) pack-a-day 10-year smoker gets a chronic respiratory illness. The number 1 is the count of the event of interest, and the number 4 is the count of the sample space. The probability, then, is the frequency with which the event or property of interest happens relative to the frequency with which the sample space event or property happens. This ratio is then expressed as a decimal (divide the event of interest number by the sample space number). In this (imaginary) case, the answer to our question is: a probability of .25. Note that this answer requires that the future will continue to take place more or less similar to the way the past has happened; if we have evidence to the contrary, then we can't use past frequencies to arrive at the probability of a future event. We will accept this important general assumption. Applying it to the smoking example, we have the assumption that future 10 year pack-a-day smokers will succumb to chronic respiratory illness at roughly the same rate as past 10 year pack-a-day smokers have. With this assumption and these statistics, it is reasonable to believe with a confidence degree of .25 that a given individual who qualifies to be in the sample space will also qualify for the event of interest. In other words, if you took a large group of pack-a-day 10-year cigarette smokers and said of each, one-by-one, "You will get a chronic respiratory illness", your prediction is true roughly one out of four times.

Let's look at another example. Suppose that you just bought a certain brand new tire for your car. What is the probability that your tire will go flat within the first week of use due to a factory defect? To find this probability you must first construct a sample space. Let's say that you had access to the manufacture's records of all the new tires sold in the past year, and this was 100,000 tires. This is now your sample space. What is the event or property of interest? It is those tires that went flat within the first week of use due to a factory defect. The records (we'll suppose) are reliable and reveal that 200 tires are in this event-of-interest category. The probability, then, is 200 out of 100,000 = 1/500 = .002. With what degree of confidence is it reasonable for you to believe that your new tire will go flat in the first week of use due to a factory defect? Using the factual probability as evidence, you should believe this statement with confidence strength .002 (which is to say that it should be very far down on your list of worries).

Summary – here are the steps in estimating and assigning factual probability:

1) Identify the sample space, defined as the set of events (or a good sample of such event if all the events can't be counted) within which the event of interest happens, and count them. This is done by consulting records, data bases, or making observations.
2) Identify the property or event of interest and count the frequency with which it happens within the sample space.
3) The probability is the ratio of the number for the frequency of the event of interest to the number for the sample space, expressed as a decimal.
4) The factual probability assigned to an event or property (i.e. that it will happen) determines the rational degree of confidence a believer should have about the truth of the statement describing that event or property.

Now we'll look at assigning pure probability, also called "a priori probability" or "classical probability". For **pure probability**, we count possibilities, not the relative frequency of actual past events. Here is an example. Suppose you wanted to know the probability of picking the ace of spades from a full deck of cards that has been completely shuffled. Note again that there are two important things. First, there is the full set of possibilities. A complete deck of cards contains 52 possibilities that you can pick from. They are equally possible picks because the deck has been completely shuffled. These possibilities make up the sample space. Second, there is the event you are interested in, namely, picking the ace of spades. There is only 1 ace of spades in a full deck of cards, so there is only one possibility that it will be picked. As in the case of factual probability, pure probability is a ratio of two numbers: the number of possibilities for the event or property of interest to happen relative to a base number of total possibilities in the sample space. Here, in the case of picking the ace of spades, we have 1/52, expressed as a decimal .02

(rounded off). As in the case of factual probability, there is an important assumption – actually two assumptions – to make. First, that the sample space count has covered *all* possibilities; in this case, a perfectly complete deck of 52 possible cards to pick from. Second, that the sample space events or property are *equally* possible; in this case, the deck is thought of as perfectly shuffled into a random distribution of cards so that each card has the same chance (1 out of 52) of being picked. Given these two assumptions of completeness and equi-possibility, and given these counts, it is reasonable to believe with a .02 degree of confidence that the ace of spades is picked.

Here is another example of assigning probability that is done by counting pure possibilities. Suppose that 8 people are sitting around a circular table, evenly spaced apart. In the middle of the table there is a spinning arrow that will randomly stop, pointing to one of the people. The lucky person the arrow stops at will receive $1000 prize. Mary is one of the 8, and she believes that she is going to win the $1000 prize. If Mary is rational, what is the degree of confidence of her belief? The sample space of possibilities = 8. They are equally possible winners. Mary is 1 out of 8. Thus, the pure probability is .125. This, then, should be her degree of confidence that she will win the $1000 prize (and, of course, she should disbelieve that she will win with a .875 strength).

Here is a third example. You roll a pair of fair dice. With what degree of confidence should you believe that a total of 7 dots will face up? Each die has 6 faces and each face has equal probability of landing up (because the dice are fair). So, there are 36 possible combinations that are equally possible. This is your sample space. Now, count the possible ways the event of interest can happen. Let's call one die A and the second one B. A can land 1 and B 6, or A can land 6 and B 1. A can land 2 and B 5, or the other way around. Finally, A = 3 and B = 4, or the other way around. No other possibilities than these 6 for the event of interest, right? So, 6/36 = 1/6 = .17. Your reasonable degree of confidence believing that 7 dots will land up should be .17.

Summary – the steps in calculating pure probability are these:

1) Identify the sample space, defined as the complete set of equal possibilities within which the event or property of interest is possible, and count these possibilities.
2) Identify the possibility that is the property or event of interest and count these within the sample space.
3) The probability is the ratio of the number for the possible events of interest to the number for the possibilities in sample space, expressed as a decimal.
4) The pure probability assigned to an event or property (i.e. that it will happen) determines the rational degree of confidence a believer should have about the truth of the statement describing that event or property.

5.4.2 Potential errors to avoid in estimating initial factual or pure probabilities.

Assigning accurate initial probabilities depends on a correct count of the items that make up the sample space and well as a correct count of the frequency of the event of interest. Sometimes these counts can easily be made, but often we must rely on our ability to estimate what these counts would be, for we are not able actually to perform a careful count (due perhaps to limited time or lack of resources or inability to access data, or even due to the relative low importance we give the decision). Whatever the reasons requiring us to estimate, it is well-known that when estimating probabilities we can go wrong in our judgments if we hold certain mistaken beliefs about probability. Let's look at some of these potential errors about probability; this will help us be on guard to avoid assigning inaccurate initial probabilities to states-of-the-world.

1) Sometimes it is easy to misunderstand exactly what a statement of probability is claiming. This might happen because a statement of probability is not worded carefully and as a result it can be confusing and subject to misinterpretation as to what exactly the probability attaches. Being clear about the meaning of probability statements will help you clearly express your own probability estimates. Suppose, for example, that you are listening to a

Boston, MA weather forecast for tomorrow and this is what you hear: The probability of rain in Boston tomorrow is 50%. What does this mean? Someone might think it means that it will rain over half of Boston tomorrow, but the experts are uncertain exactly where. Others might think it means that it will rain for half the day tomorrow over all of Boston, but the experts are not sure about the exact times. It might not rain all day in Boston tomorrow, and it might not rain everywhere in Boston tomorrow. But this is not what is being claimed in this forecast. The event of interest is rain in Boston tomorrow, it is not times of day of rain or exact locations of rain. What would the sample space be? It would be all the past days in Boston having similar variables of temperature, season, weather conditions, etc. as today, for which the forecasters have records. Given this data, the forecasters count how many days in this sample space had rain the next day. They find out, we'll suppose, that one half of the sample space days did and the other half didn't. Projecting these numbers from the past onto today and tomorrow in Boston gives us the forecast. (Note again the need to use the important assumption – the future will be similar to the past – in assigning factual probabilities such as this.)

2) The second easily-made mistake has to do with the nature of memory. It is pretty well established by empirical studies that repetition strengthens memory. This is also common knowledge. If we want to remember something, a common method we use is to repeat it to ourselves over-and-over. So the rule is: frequent experience establishes and strengthens memory. Of course, other things can also establish a strong memory. A traumatic experience might leave a person with a life-long vivid memory, even though it happened only once. Because frequent experience and repetition are not the only causes of strong memory, the rule does not work in reverse. While it is true that frequency strengthens memory, it is not true that if a person has a strong memory of an experience and can easily recall it, then the experience being remembered happened frequently. Yet people sometimes believe this reversed rule. This is called the **availability error.** The error is to think that if something easily comes to mind or is easily available to memory then it must be highly probable

because it must (have) happen frequently. Here is an example where this error can easily lead to assigning inaccurate initial probabilities. Let assume for the moment that trains and planes have about the same safety record. (I don't know this for a fact, but it seems a not unreasonable assumption.) Thus, the probability of injury or death in a train wreck and the probability of injury or death in a plane crash are roughly equal, and both are very low, much lower than the probability of injury or death in an automobile accident. Yet in the aftermath of the September $11^{th}$ 2001 terrorist attack on the WTC in New York City, these plane crashes were very available to many people's memory. They were vivid and horrifying. There were wide spread reports that people decided to drive or take the train instead of flying because they were afraid of dying in a plane crash. And yet the probability of dying in an automobile accident was, and still is, much greater than dying in a plane crash, and death by plane remained roughly equal (given our assumption) to death in a train wreck. The availability error seems to be at work here, leading to exaggerated probability estimates, and therefore fears, of plane crashes.

3) If asked to estimate the factual probability of some event, one strategy someone might resort to is imagination, especially in the absence of statistical data. Say someone asks you what the probability is of having a traffic accident while driving down the main street in your town. You know the main street very well, and feel that this is information you should be able to provide the questioner. You realize that you do not know any actual statistics. Instead, in imagination you picture yourself driving down your familiar main street, and you try to estimate the chances of getting into a traffic accident. You recognize that there are over ten intersections on the main street and each one could, in your mind, easily lead to an imagined traffic accident. So, you answer that the probability is pretty good of a traffic accident while driving down the main street. You may sound knowledgeable to someone who does not know the facts, but your probability estimate is almost certainly way off base. Substituting imagination for actual statistics is called **scenario thinking**. An agent will run through a scenario in imagination hoping to read off the frequencies from the imagined events. Why is

scenario thinking bound to fail as a method of estimating probabilities? It fails because an agent cannot extract information from an imaginary scene that does not contain the information in the first place. If the agent truly does not know the relevant frequencies, the statistics, then the agent cannot introduce them into the imagined scenario. If they are not contained in the imagined scenario, the agent cannot extract them from the imagined scenario. Scenario thinking will give an agent only imaginary probabilities, not factual probabilities.

4) Suppose you are told that a high percentage of firefighters admitted being fascinated with fire as a teenager. Let's say a district of a large city employs 20 firefighters and 80 police officers. Relatively few police officers admitted that they were fascinated with fire as a teenager, say only 20%. But a large number of firefighters admitted to this, say 80%. You are now introduced to someone employed by the district and are told that this person admitted to being fascinated with fire as a teenager, and that the person is either a police officer or a firefighter. You are asked to guess which one is more probable. How would you answer? Many people would say that it is more probable that the person is a firefighter. But this is false. It is a case of **ignoring the base rate frequency**. In fact, it is equally probably that the person is a police officer. While the relative number of police officer fascinated by fire as a teen is low, notice that there are many more police officers employed by the district than fire fighters. This is called the base rate frequency. Ignoring the base rate frequency leads to the error of assigning a higher probability that the person in question is a firefighter. Here is another case. Suppose you knew that stockbrokers loved to read the financial reports that companies put out, but that it is very rare for people who are not stockbrokers read such financial reports. You are introduced to someone who tells you that she loves to read financial reports. What is more probable, that she is or is not a stockbroker? If you ignore the base rate frequency, you'll incorrectly think that it is more probable she is a stockbroker. The base rate frequency of people who are not stockbrokers is vastly greater than the frequency of stockbrokers. As a result, the number of the former group who love to read financial

reports, even though rare, is bound to be much larger than the number of the latter group who love to read such reports. Thus, it is far more probable that the person you are introduced to who loves to read financial reports is not a stockbroker than it is that she is a stockbroker.

5) Many people confuse the idea of average and the idea of frequency. If the average family has two children, should we believe that families with two children occur frequently and thus are more probable than, say, families having five children? If the average income in a town is, say, $60,000, does this mean that $60,000 is the most frequently occurring income in the town? In each case the answer is no. To try to estimate probabilities by appealing to what is average is called **the error of the raw mean**. Take a range of values: they could be incomes in a town, or temperatures every hour in a day or at noon every day in a week or a year, or the number of children in families that own farms,…, whatever. As an example, lets suppose that temperatures are taken at noon for one week. Mon. = 63 degrees F, Tues. = 68, Wed. = 76, Thur. = 90, Fri. = 81 , Sat. = 84, Sun. = 84. The average temperature is the raw mean of this range of values. You find this by adding these seven temperature values and dividing by 7. The average noontime temperature for this week is 78 degrees. Now notice how frequently 78 degrees F occurs – not at all. Clearly, "average" can't mean the same thing as "most frequently occurring value", and thus it is an error to believe that the closer to the average a value is the more probable it is. What is average might be probable, but clearly it need not be so. Instead of "average", the ideas of "ordinary", "common", "easy", or "typical" are more closely linked to the idea of "frequently occurring", and might sometimes provide a guide for estimating probability whereas "average" cannot.

6) Just as people sometimes confuse average with frequently occurring, there is a related error in confusing unique with infrequent. Thus, many people believe that the more unique something is the more improbable it is. Let's call this **the uniqueness error**. Here is a counterexample to this belief. Let's say that you flip a fair coin three times. It would be

unique if the coin came up heads three times in a row, more unique someone might think than if it came up some mixture of heads and tails, say, THT. Thus, someone might be lead to believe that HHH is less probable (more improbable) than THT. But this belief is false. Getting HHH in three flips is just as probable as getting THT; each has a probability of .125. Just as it is easy to make the error of the raw mean by confusing “average” with “frequent,” in the uniqueness error it seems that “unique” is easily confused with the idea of “infrequent.” Being unique does not necessarily mean a thing or event is highly improbable.

5.4.3 Combining initial probabilities.

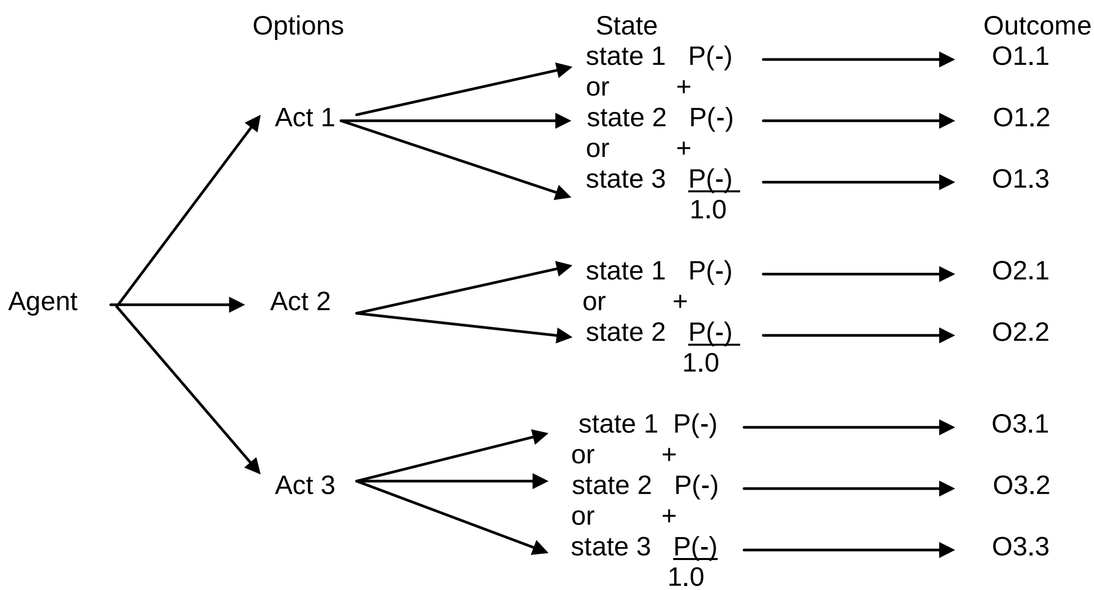
Once initial factual or pure probabilities have been assigned to states-of-the-world, these probabilities can be combined according to some basic rules of the probability calculus. For our purposes, 4 rules are needed. As we review these basic rules, examples will be offered to show how they are used in setting up the framework for decisions under risk.

Rule 1. All probabilities are normalized. Information is “normalized”, you will recall, if it is represented on a scale that sums to 1.0. The scale for probability is: 0 – 1. A state-of-the-world that cannot fail to happen is assigned the value 1. Thus, 1 represents the agent’s maximum degree of confidence – certainty. A state-of-the-world that cannot happen is assigned the value 0. Thus, 0 represents the agent’s minimum degree of confidence – disbelief. Thus, all degrees of confidence, are represented by numbers between 0 and 1. If we let “P” represent the probability and “a” represent a state-of-the-world, then we can symbolize rule 1 as: $0 \leq P(a) \leq 1$

Rule 2. Two states-of-the-world, a and b, that are mutually exclusive (meaning only one can happen, not both) and jointly exhaustive (meaning there is no third state that can possibly happen) must have probabilities that sum to 1. Thus, $P(a) + P(b) = 1$, and $P(a) = 1$ minus $P(b)$, and $P(b) = 1$ minus $P(a)$.

Rule 3. The disjunction (addition) rule: For any two states-of-the-world, a and b, that are exclusive: P(a or b) = P(a) + P(b).

Rules 1, 2 and 3 apply to the possible states-of-the-world for any <u>one option</u>. Say an option has three possible states, each yielding a different outcome. The agent must assign initial factual or pure probabilities to each of these three states in such a way that rules 1, 2 and 3 are obeyed. The states are alternatives and so are connected by “or”. Their probabilities must sum to 1. In effect, the agent must include all states possible for an option under risk until the agent is certain that no alternative state has been left out (each one yielding a different outcome). If we put this into our framework we get this structure for an agent with three options, two of which have three possible states and one (Act 2) having two possible states.




In this decision diagram, each option has states as possible alternative conditions that lead to different outcomes. The agent must be certain that one or the other state will be in place for each option. What the agent is uncertain about is which state. The probability of each state happening (actual numbers has been left out of this diagram and little “-“ ’s put in their place) will provide the agent with the evidence to form a reasonable degree of confidence about each outcome for each act. Note that the number of alternative states for decisions under risk must be more than one, but there is no upper limit on how many there might be. It is the agent's responsibility to discover and list all alternative states within each option, if a rational choice is to be made.

Before going on to the next rule, here are two examples where you should apply rules 1, 2 and 3. Suppose you are a reporter who must quickly leave a dangerous war zone. You must fly over a high mountain range for many hours until you reach safety. All planes have departed except two old small prop planes. You have a choice. One plane has one engine, and the other plane has two engines. You are told by a good source that all three engines are in equal mechanical condition and that the reliability of these engines is that they fail once in every 100,000 trials. You are also told that the planes are in equal mechanical condition, the pilots equally skilled, and that the two-engine plane cannot fly with one engine. You must now quickly decide which plane to take to escape the danger. Which plane has the greater probability of yielding the outcome of your safety? Answer: The plane with one engine. Why: It has a 1 in 100,000 probability of crashing. But the plane with two engines will go down if either engine #1 *or* engine #2 fails. So, by rule 3, P(#1 or #2) failing = P(#1) + P(#2) = 1/100,000 + 1/100,000 = 2/100,000 = 1/50,000. Thus, in this example you are twice as much at risk of dying due to engine failure in the plane with two engines than you are in the plane with one engine.

We considered above an example of assigning an initial factual probability to a new tire going flat due to a manufactures defect. In our imaginary example, we estimated the probability to be .002. Suppose that you have just bought a set of four such tires and they are now mounted on the four wheels of your car. What is the probability of getting a flat in one of the tires within the first week of use due to a defect? Well, a flat could happen in new tire #1 or in #2 or in #3 or in #4. By rule 3, the "or's" become "+'s". Tire #1 P(.002) + #2 P(.002) + #3 P(.002) + #4 P(.002) = P(.008). Thus, it is reasonable to believe that you will get a flat (in this imaginary example) with a degree of confidence .008.

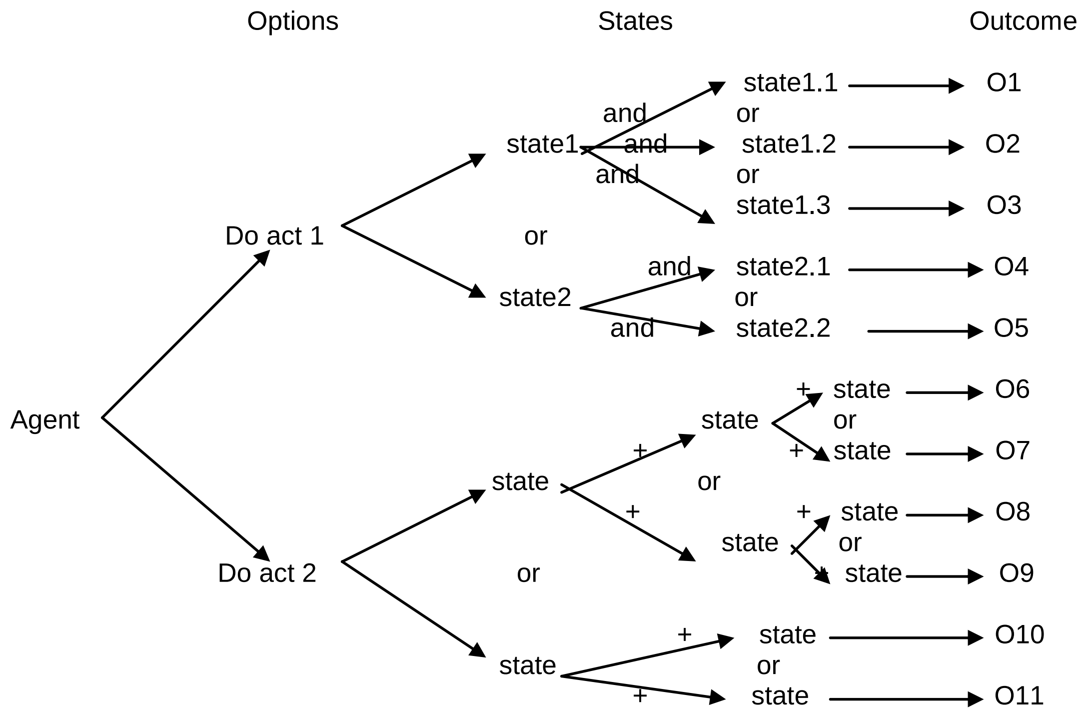
Rule 4. The conjunction (multiplication) rule: For any two states-of-the-world, a and b, that are independent, P(a and b) = P(a) x P(b). Two states are independent if the probability of either one does not depend on the probability of the other. For example, what is the probability of getting heads two times in a row flipping a fair coin? Clearly, getting heads on one of the flips has no affect on getting heads on the other flip. So, P(heads and heads) = P(heads) x P(heads) = P(.5) x P(.5) = P (.25). Here is another example. What is the probability of picking an ace from a full deck of perfectly shuffled cards <u>and</u> flipping tails with a fair coin? Again, these are completely independent events. The probability of drawing an ace is 4/52 = 1/13 and the probability of getting tails is 1/2. By rule 4: 1/13 x 1/2 = 1/26 = .04.

If the states are not independent – that is, if the probability of one affects the probability of the other – then a variation of rule 4 must be used. For any two states, a and b, such that b depends upon a, P(a and b) = P(a) x P(b/a). Read the second part of this rule as saying: the probability of state b given that state a happens. For example, what is the probability of a college student graduating college and getting a job (as a college graduate) with a starting salary over $50,000? Clearly, the probability of getting a job with a starting salary of over $50,000 depends very much on whether a person is or is not a college graduate. Let's work this out with some imaginary numbers. Suppose 4 out of 5 college students graduate college. That would mean that 1 in 5 college students, for whatever reason, never graduate college. Let's further suppose that only 1 in 10,000 people who do not graduate college get jobs with starting salaries over $50,000, and the rest (9,999) get jobs with starting salaries under $50,000. But suppose that 1 in 100 college graduates get jobs with starting salaries over $50,000, while 99 do not. Note that there are different probabilities of getting such a job, depending on whether a person graduates college or not. So, the probability of a college student both graduating college <u>and</u> getting such a job (as a college graduate) must be calculated using the dependent version of rule 4. P(4/5) x P(1/100) = 4/500 = 1/125 = .008.

Here is another example that serves to contrast the difference between the independent and the dependent versions of rule 4. Suppose you are asked to pick a card from a perfectly shuffled full deck of cards. What is the probability of picking the ace of spades? 1 in 52. What is the probability of picking the ace of spades on the second try if the first pick was not the ace of spades? Well, that depends on whether the first pick was randomly put back into the deck or not. If it was, the second try is independent of the first try and the probability is again 1/52. But if the first pick is not put back, then the probability of picking the ace of spades on the second try depends on the first try, for on the second try the deck now has one less card to worry about. The probability would now be 1/51.

Rule 4 applies to decisions under risk when there are multi-stage states-of-the-world. Sometimes an act will yield the intended outcome if a single state-of-the-world is in place. For example, suppose it is getting dark in your dorm room and you need light to continue studying for a test. There are two light sources, a desk lamp and a ceiling light. You decide to put on the ceiling light, but there is a slight chance that the old burnt out bulb has not yet been replaced by the maintenance staff. Its replacement is the single state that must exist if your action is to yield the desired outcome of sufficient light from the ceiling light to continue studying. In most decision situations, however, a whole sequence of states must exist if the action decided on is to yield the desired outcome. Look at someone who decides to get a college degree. A huge number of states-of-the-world must be in place, semester after semester, if this decision is to yield the outcome. A **multi-stage state decision** means that more than one state-of-the-world must exist if an option is to yield a given outcome. Here is another example. Suppose that you are visiting a friend in an unfamiliar city. One night an emergency comes up and you must drive to the pharmacy for medicine for your friend. As you drive you notice that her car is almost out of fuel. Also, you are not sure if the pharmacy you were directed to is still open. In order for the desired outcome to happen, the world must “cooperate” in two ways with your decision to get the medication. First, the car can't run out of fuel (or an open fuel station has to be available). Second, the pharmacy must not have closed for the night. This is a case of multi-stage states. As you can see, these two states are independent; the probability of one (having enough fuel) does not affect the probability of the other (pharmacy still open).

For the case of a multi-stage decision with two options, the abstract diagram looks like this.




In this hypothetical decision structure, note that there are 11 outcomes, depending on which sequence of multi-stage states exists. Wherever the states are related by “or”, rule 3 applies and the probabilities of the states must sum to 1. Wherever the states are connected by “and” (for example, outcome 5 requires state 2 *and* state 2.2 to exist), rule 4 applies and the agent’s degree of confidence is calculated by multiplying the probabilities of the required states. Note that for option 1 there are two-stage states that are required for the outcomes, but that option 2 is an example of three-stage states. Multi-stage states means that more than one state is required for the outcome to be produced, there is no limit to the number of states that might be needed. For each stage, however, the agent must try to discover and list all the alternative states that might compete with the desired one, and the probabilities of these must sum to 1, per rules 1, 2 and 3.

Shortly we will be applying these ideas to solve decision problems under risk. But before doing so, let's look at two very common errors people make when combining probabilities.

#### 5.4.3.1 Potential errors to avoid in combining initial factual or pure probabilities

Just as it is important to assign accurate initial probabilities if risky decision making is to be done rationally, it is likewise important to combine initial probabilities correctly. There are two easily made errors that can lead us astray when combining probabilities: (1) the gambler's fallacy and (2) the equi-probability fallacy. Because it is so easy to make these errors, they are widely found. Describing them will help us be on guard to avoid making these errors when combining probabilities.

(1) The gambler's fallacy

(i) The story is told that a roulette wheel in one of the gambling casinos in Monte Carlo once came up red, astonishingly, 26 times in a row. Because the wheel was not biased toward red, each spin had an equal chance of the ball landing on red or black. Yet it landed on red over and over. As red came up again and again, the table increasingly drew the attention of the gamblers who were betting in the casino that night. They reasoned that with so many reds in a row, black is due to come up because the wheel was unbiased and would be correcting itself to restore the even balance between red and black. But with each new spin red came up, and the crowd became more excited in the expectation that black gained a stronger and stronger chance of coming up. The crowds around the table bet larger and larger sums on black. The more they lost because the ball kept landing on red, the more convinced they became that black was now the sure bet. Many people lost large amounts of money, and some lost all they had. Black did not come up until the 27$^{th}$ spin. The error these gamblers were making has come to be called the "Monte Carlo Fallacy" and more commonly the "Gamblers Fallacy". Here are some other versions of it that you may recognize.

(ii) A person in a casino is losing at a slot machine, losing over and over. Yet she believes that her luck is due to change the very next try precisely because she has lost so many times. And so she stays on the machine, and keeps losing.

(iii) A person gets lucky and wins something with a lottery ticket. When he buys another ticket he does not use the same set of numbers he used for the ticket he won with, reasoning that it is highly improbable that the same set of number would win for him twice.

(iv) A person knows that automobile accidents are common but that train crashes are rare. There has just been a train crash. So she decides to travel by train, reasoning that it is even more improbable than usual that there will be a train crash soon, because there has just been one.

(v) A person tells his friend that he drew the ace of spades from a well-shuffled deck of cards. The friend responds that that is a very unlikely event and so he must have been drawing and reshuffling cards for a long time before the ace of spades came up.

These are all cases of the **gambler's fallacy**. What is the error being made? In each case there is a series of similar events. Spins of a roulette wheel, tries at a slot machine, selecting lottery ticket numbers, train crashes, picking cards from a deck. Even though the events in each of these series are <u>similar</u>, it is important to see that the events in each series are <u>independent</u> of one another. And yet in each case someone estimated the probability of one event in the series as if it was influenced by the earlier events in the series. It is an error to treat independent events as if they were dependent; "similar events" does not mean "dependent events." If you flip a fair coin and it comes up heads 5 times in a row, on the $6^{th}$ flip tails is still a 50-50 chance. Tails does not gain in probability just because the coin has landed heads 5 times in a row, as if the coin "remembers" this and now tries to "restore" the 50-50 balance with a run of tails. Roulette wheels have no memory, nor do slot machines, lottery ticket numbers, train crashes, or cards picked from the deck. These events can't influence how future events in the series will turn out, nor can the person's knowledge of the past series influence how the future events in the series take place. The reason is that in each case we have a series of independent events. If a person buys lottery tickets week after week believing that the day must soon come when he will win because he has lost so many times, this is the gambler's fallacy. Likewise, if a person who has won the lottery stops buying tickets believing that the chance of winning again is now lower than the chance of winning once, this is again the same error. (Note that the probability of winning the lottery again is not affected by having won it before. But this is not the same as asking the probability of winning the lottery two times. Here we must use rule 4 to combine two initial probabilities, and winning two times is clearly less probable than winning again after having won before.) In each case the lottery, if it is fair, does not retain information as to who has and who has not won previous drawings. There is no way for it to influence who the next winner will be. Understanding the difference between independent and dependent events will help you avoid using the wrong version of Rule 4 when combining initial probabilities.

(2) The equi-probability fallacy

If a fair coin is flipped, heads and tails have equal probability of landing up. Because there are only two possibilities, heads or tails (let's rule out the possibility that the coin lands on its edge), it would be correct to say that heads and not heads are equally probable. This is correct because "not heads" equals tails, and there are no other possibilities included in "not heads." Landing tails is the only way for the coin to land not heads. It is certain that the coin will land heads or tails, so: P(H or T) = 1. Heads must get a probability equal to not heads (tails), and since there are only two possibilities over which the value 1 must be equally distributed, heads and not heads each has a probability of .5 and P(H or not H) = 1. This reasoning is correct for the case of the coin, but it can easily be misapplied in cases to which it does not fit. The error is called the **equi-probability fallacy**. Here are some examples of this error.

(i) From a full deck of perfectly shuffled cards, each card has an equal probability of being picked. But suppose someone reasoned that from such a deck of cards there are only two possibilities – drawing an ace or drawing a card that is not an ace. Because there are only two possibilities, they must be equally probable. So, there is a 50-50 chance of picking an ace from a deck of cards.

(ii) Suppose a person thinks that his chance of dying from cancer is .5. He reasons like this: I can either die of cancer or not die of cancer; there are only these two possibilities. I must die, that's certain. So, the probability of my dying = 1. Because there are only two possibilities, dying from cancer or not dying from cancer, there is a 50-50 chance of my dying from cancer.

(iii) Many families have 3 children. The probability that a family with 3 children will have 3 girls must be 1/3. Why would someone think this? Well, there are only three possibilities: there could be 3 girls, or there could be a mixture of girls and boys, or there could be 3 boys. Given that there are 3 children and that there are just 3 possibilities for 3 children, the probability of 3 girls must be 1 out of 3.

What is the error in these examples? Look back at the case of flipping the coin. Heads and not heads are equi-probable because "not heads" has only one way to happen, namely, when tails lands up. This means that the number of ways for heads to happen equals the number of ways for not heads to happen. Thus, they are equi-probable. But in the first erroneous case, the number of ways for "not-ace" to be picked is not equal to the number of ways for an ace to be picked. There are 4 aces, and so "ace" has 4 possibilities. But there are 48 non-aces in a normal deck of cards, and so there are 48 ways for non-ace to happen when picking a card. Clearly, the probability of getting an ace is not equal to the probability of picking a non-ace card, the count of the event of interest yields very different numbers. The same analysis applies to not dying from cancer; this has many more ways to happen than the event of dying from cancer. Finally, there are many more ways to get a mixture of girls and boys (6 ways) in a family with 3 children than there is to the event of all girls or all boys. Thus, these 3 possibilities can't be equi-probable. What causes the error is a failure to count how many ways the alternative to the event of interest can happen. One possibility, for example picking a non-ace from a deck of cards, might happen in many different ways; to correctly combine initial probabilities the number of ways cannot be overlooked or lumped together under one general possibility.

We are now in a position to frame and solve decision problems under risk, the topic of the next chapter.

EXERCISE:

This chapter is rich in new concepts that are central for understanding and practicing material that will be coming. Rather than practice in assigning initial probabilities and calculating combined probabilities (which you will have gotten if you already had a basic college level general mathematics course, and which we will be practicing in the exercises in the next two chapters), the exercises for this chapter are designed to reinforce your grasp of concepts.

1) Provide (i) a definition (in your own words!) and (ii) an example of each of these concepts. For any that you are not clear about, try to avoid consulting the Glossary and instead re-read the appropriate section in this chapter.

a. default certainty
b. direct evidence certainty
c. risk
d. positive outcome

m. sample space
n. availability error
o. scenario thinking
p. ignoring base rate error

e. negative outcome

f. hope limit

g. security limit

h. belief and disbelief (as propositional attitudes)

i. degree of confidence

j. reasonable degree of confidence

k. factual vs. pure probability

l. event of interest

q. error of raw mean

r. uniqueness error

s. disjunction rule (add)

t. conjunction rule (multiply)

u. single vs. multi-stage states

v. gambler's fallacy

w. equi-probability error

2) For each of the following situations, say if the person is:

a. Someone who has an unreasonable degree of confidence

b. Someone who has an excessive hope limit

c. Someone who has an excessive security limit.

Explain your answer. What in the situation makes the person unreasonable or excessive? Try to narrow each answer to just one of the above, but if you think more than one apply, say why.

(i) Mary has met John only once at a party. Conversation with John was easily, and they spent most of the evening talking, finding that they have many things in common. He seemed intelligent, interesting, and dynamic; Mary thought about him weeks after the party. One day several weeks later, to her surprise and delight, John calls Mary with an exciting offer, a chance to buy into a time-sharing apartment together at their favorite beach. He says that he's been thinking of her and when this time-sharing opportunity came up he wanted to share it with her. However, John is between bonuses at this point and can't come up with the money. He asks Mary to trust him, and to agree to send him the $3000 payment. She believes John and sends him the money.

(ii) Harry must undergo surgery, and must decide between two widely used procedures: one involves lasers, and the other doesn't. The doctors tell him that each procedure has 3 possible outcomes. (1) The most likely outcome is that the surgery will cure the problem completely. (2) There is a slight chance the surgery will temporarily cure the problem and it will return within 3-to-5 years, in which case he will have to repeat the surgery. (3) In very rare cases the surgery will not work, in which case Harry will have to undergo the same surgery within a week. As Harry tries to decide which procedure to choose, he finds that he is becoming increasingly fearful that the outcome in his case will be death.

(iii) Sue has a tough semester coming up. She must take several of her math and science requirements next semester, and they are known to be "killer" courses. In the organic chemistry course, it is the rare student who earns higher than B grade. Sue has a super grade point average, 3.75 out of 4, primarily (she knows) because she has not yet had to take these hard courses. In addition, from her high school experience she believes that she's "not good" at math and science. Instead of preparing to see her GPA go down, Sue hopes her GPA goes up next semester.

(iv) A video rental company has recently decided to open a branch in a local mall. They are sure that the branch will do well. The spot, however, is at the far end of the mall that gets minimal traffic. In addition, several small businesses have tried to make it in the same spot and have had to go out of business. On top of that, the in-store video rental market is rapidly declining due to on-line video rental services. Yet, the company continues to go ahead with its plans for a new branch there, feeling certain that it will succeed.

(v) Smith is selling her used car. She looks up its value, given the condition it's in, and finds that her model sells for a low of $1000 and a high of $1600. Smith, however, puts "$2000 firm" in the classified add for her car.

(vi)  Jones has taken his friend to a fancy restaurant, hoping to impress her. He is surprised at the prices; the cheapest offer, without wine, is $75 per person and the most expensive, again without wine, is $400 per person. He decides on the least expensive offer, but as he peruses the menu he worries that the bill could set him back over $600 not counting any wine they order.

<u>Sources and Suggested Readings</u>:

This chapter draws on material from: Skyrms (2000) Chapters VI and VII, Hacking (2001) Chapters 11-15, Resnik (1987) Chapter 3, and Mullen and Roth (2002) Chapter 4.  These authors are highly recommended both for presentations of probability and for the connection between belief and probability. Skyrms was particularly relied on in presenting objective relative frequency as evidence for forming reasonable subjective belief strengths. Of course, any student who has had a basic general mathematics course will already be familiar with probability (at least the objective or relative frequency concept) and the rules for combining probabilities. For belief in the context of decisions Jeffery's (1983) Chapter 4 is both a clear and a seminal source. The classic work on our vulnerability to error in probability judgments is done by Kahneman and Tversky; their article "Judgements under uncertainty: heuristics and biases" is a must (in Moser (1990) Chapter 7). Part III of Holyoak and Morrison contains an excellent summary of the literature on the ways probability judgments can go wrong. Plous (1993) is a good source for experiments that build on the work of Kahneman and Tversky concerning the pitfalls of judging probability.