

Social Studies 201**October 18-20, 2004****Probability**

See text, chapter 6, sections 6.1 and 6.2

1. Introduction

In the first part of this semester, we studied descriptive statistics in the form of distributions, histograms, and measures of central tendency and variation. Part I of the text is devoted to these topics.

For the remainder of the semester, we will be examining probability and inferential statistics – the topics of Part II of the text. Chapter 6 of the text is devoted to definitions and examples of probability and the normal distribution. (We will not be studying the binomial distribution of sections 6.3 and 6.5 this semester). The topics of probability and the normal distribution differ from the descriptive statistics of the first part of the semester, but set the stage for the inferential statistics of chapters 7-11. These notes give an introduction to the study of probability.

Reasons for studying probability

In this course, we study probability as a way of illustrating, understanding, and using a statistical approach to social analysis. That is, we do not study probability as a subject of interest in itself, and do not develop probability theory, but use only a few principles and methods of probability. In the following notes, there is a discussion of the principles of probability and the three different interpretations of probability – classical or theoretical, frequency or empirical, and subjective. While games of chance are used to illustrate basic principles of probability, we will not spend much time considering complex probabilities associated with such games. Rather, we will apply basic principles of probability to survey sampling examples. We will use these to illustrate interpretations of cross-classification tables and the ideas of independence or dependence of variables. Following this, we will discuss the normal distribution.

Probability is emphasized in statistics and in modern society for three main reasons. The first two reasons are primarily statistical and the third relates to the structure of modern society.

- **Inferential statistics.** Probability is required for inferential statistics, that is, using a sample or experiment to estimate characteristics of, and conduct hypothesis tests about, a population. Much data concerning social issues emerges from samples or experiments, where a relatively small number of members of a population is studied or surveyed. From this small group, a researcher makes inferences about characteristics of a whole population. When doing this, it is common to select subjects for the sample survey or experiment using principles of probability. Where this is the case, inferences about the whole population can be drawn from the data obtained from sample or experiment. But such inferences always have a probability attached to them. That is, conclusions drawn from samples are not definitive statements about a population, but statements that generally have a high probability of being correct. There is always a chance that the statements are incorrect – hopefully this chance is small.

Example – Opinion polls.

Obtain statement of a confidence interval from an opinion poll.

- **Models.** Many explanations of social phenomena are probabilistic, in that they do not provide definitive statements about a population but have a probability of being correct attached to them. That is, rather than being deterministic models stating that A causes B, social explanations from most social science models have probabilities attached to them.

A deterministic model can be illustrated as follows:

$$A \rightarrow B$$

while a probabilistic model would have a probability attached to it.

$$A \rightarrow B \text{ with probability } p, \text{ where } 0 \leq p \leq 1.$$

An example of a deterministic model is that of gravity. If I shove a book off a table (A), then that book is sure to fall to the floor (B).

As an example of a probabilistic model in the social sciences, consider the relationship between an individual's education and the same individual's income twenty years after completing that education. While surveys and socioeconomic models have demonstrated that more education (A) is associated with higher income (B), there is no assurance that any specific individual who obtains higher education will be assured of higher income. What the researcher can say is that on average, or generally, there is a greater probability or chance that an individual with more education will have higher income than an individual with less education. That is, the social science model predicts that more education leads to higher income, but with a particular probability associated with this prediction.

The above examples also illustrate one of the limitations of probabilistic models – they do not necessarily predict individual behaviour or results. At the same time, these probabilistic models may be good at predicting social behaviour for a group, or in the aggregate. For example, small samples that are well constructed can reasonably accurately predict election results, but they cannot predict how a specific individual votes.

While we will not study many models in this class, the normal or bell curve is one example of a statistical model that is widely applied to social situations. For example, instructors may assign grades by matching them with the bell curve. Or it may be found that the distribution of grades is close to being normally distributed, even if that was not intended by the instructor – you can examine the distribution of GPA in the SSAE98 data set, with the normal distribution superimposed, to see that the distribution is close to normal.

- **Quantifying uncertainty.** For much of history, there was great uncertainty concerning what occurred in the natural and social world – events happened and these might have been explained by referring to the gods or spirits. In the modern world, natural and social sciences have attempted to develop explanations for events – many of these explanations are sound and can be used to predict outcomes of social situations. While some of these explanations are deterministic, many are probabilistic. Assigning numerical probabilities to events is a means of quantifying uncertainty, thus allowing some predictability,

and possibly control, over events and outcomes.

While the development of science has allowed quantification of many uncertainties, the expansion of business and administration has been the major impetus for better planning, prediction, and control over uncertainties. Life insurance companies were among the first to develop statistics of death rates, in an attempt to better predict these rates. This allowed these companies to quantify their potential losses and set premiums to cover these losses plus return a profit to the owners. Large corporations and administrators in institutions such as governments, universities, hospitals, and schools attempt to predict future events. They often use probabilities and probabilistic models to assist in such explanation and prediction.

Terms for probability. Each of the following terms is used when referring to probability in this course. While there are differences in meaning among these terms, for this course each is synonymous with probability.

- probability
- likelihood
- odds
- chances
- risk

Other terms such as **propotion** and **incidence** are similar to probability, in that they refer to the number of times an event occurs per member of a population. Two concepts closely connected with probability are **random**, as in random sample, and **stochastic**, a mathematical term sometimes used to describe probabilistic processes.

Ideal games of chance or probability experiments

Principles of probability have often been illustrated using games of chance. Some well known games are as follows.

- Flipping a coin.
- Rolling a single die with six sides, or a pair of dice.
- Drawing a card from a deck of fifty-two cards.
- Spinning a roulette wheel.

One trial of any of the above can be considered to be an ideal game of chance or, what I refer to as an ideal probability experiment. These games or experiments have characteristics that make them suitable for illustrating and determining probabilities.

Characteristics of ideal probability experiments

In a probability experiment or game of chance, it is not possible to predict an exact outcome. However, there is a certain regularity or pattern to outcomes over a series of trials of the game or experiment. For example, repeated flipping of a coin results in approximately one-half heads and one-half tails. These ideal experiments or games of chance have four characteristics that result in the regular patterns observed over the set of outcomes. These characteristics are as follows – the experiment of flipping a coin, with head or tail as the two possible outcomes, is used as an example.

1. **Exact outcome unknown.** The exact outcome of a probability experiment is not known prior to conducting the experiment. Before flipping a coin, you do not know if the outcome will be a head or a tail. If you can predict whether it will be a head or tail, I'd suspect you of tampering with the coin or the flipping procedure; or I'd think you have powers of clairvoyance.
2. **All possible outcomes are known.** In a probability experiment, you must be able to list, or be able to imagine, the set of all possible outcomes. Before flipping a coin, you know that either a head or a tail must be the outcome – that is, the outcome of flipping a coin cannot be a heart (drawing from a deck of cards) or obtaining a four (as in

rolling a die). Knowing all the possibilities allows you to organize or classify outcomes, and attach a probability to each outcome.

3. **Equally likely outcomes.** In most probability experiments, each possible outcome is as likely to occur as any other outcome. Or, the structure of the game is such that no particular outcome is favoured or discriminated against. For a coin, a head is as likely to occur as a tail, or a tail is as likely to occur as a head. Neither is more favoured over the other in procedures associated with the structure of the game – it is the flipping of the two-sided coin that make a head or tail equally likely to occur.
4. **Repeatable under uniform conditions.** Probability experiments can be performed repeatedly under exactly the same conditions each time. A coin can be flipped hundreds or thousands of times and each flip is conducted under exactly the same conditions. The coin is constructed from a hard metal, so the coin does not wear out; the flip of the coin produces the same structural procedure each time; and the way the coin lands produces either a head or a tail.

Regularities or patterns. The above four conditions produce a regularity or a pattern to the outcomes. This makes it possible to predict the pattern of outcomes, or the average number of times a specific outcome occurs. If you flip the coin many times, you expect a head to result from about one-half of the flips and tails to result from the other one-half of the flips, give or take a few on either side of exactly one-half. While it is not possible to predict whether head or tail occur on any one flip of the coin, heads occur about one time in two, and tails about one time in two.

As a result of this regularity or set of patterns, probabilities are meaningful and can be calculated. When flipping a coin, the probabilities are one-half for tails and one-half for heads. In the case of other probability experiments that conform to the above criteria, the chances associated with each outcome can also be determined.

Approaches to probability

There are three approaches to probability – theoretical or classical, frequency, and subjective or judgment. Theoretical probabilities (Section A) can be determined on the basis of reasoning alone; probabilities obtained

from the frequency interpretation (Section B) require obtaining data about the incidence or likelihood of particular outcomes; and the subjective or judgment interpretation of probability (Section C) is a best guess about the chance of a particular outcome.

Each of the three approaches to probability provides a distinct way of defining and thinking about probabilities. The main features of each approach are outlined in the follow sections.

A. Theoretical or classical approach

This is the approach associated with situations conforming to the principles of the ideal games of chance or probability experiments discussed above. This approach is termed classical in that it is one way that mathematicians developed probability theory. It is theoretical in that certain probabilities can be determined entirely on the basis of theoretical reasoning, without reference to actual data or experience.

The classical approach to probability provides a means of defining and determining probabilities as follows.

Definition – probability of an outcome. If a probability experiment has N possible outcomes, where each of these outcomes is equally likely, then each outcome has a probability $1/N$ of occurring.

Examples of situations where the classical approach can be used to obtain probabilities are the traditional games of chance – flipping a coin, rolling a die, or drawing a card from a well shuffled deck of cards. In certain conditions, random sampling also conforms to these principles. Examples of how the definition can be applied in these situations are as follows.

Example 6.1 – Probabilities for flipping a coin. If a coin is flipped, there are two possible outcome, head (H) or tail (T). Each of these two outcomes is equally likely and the method of flipping a coin means that the principles associated with ideal probability experiments apply. There are only two outcomes, H and T, so $N = 2$. From the definition, when a coin is flipped once the probability of a head is $1/N = 1/2 = 0.5$. Similarly, for a single flip of a coin, the probability of a tail occurring is $1/N = 1/2 = 0.5$.

Example 6.2 – Probabilities for rolling a single die. When you roll a single six-sided die, there are six possible outcomes: 1, 2, 3, 4, 5, or 6. Each of these six outcomes is equally likely and the rolling of the die means that the principles associated with ideal games of chance apply. Given that there are six outcomes, $N = 6$. Since each of these outcomes is equally likely, the probability for each particular outcome is $1/N = 1/6$. The probability of obtaining a 1 on a roll of a single die is $1/6$, the probability of obtaining a 2 is $1/6$, and so on.

The list of all the possible outcomes for rolling a single die, along with the associated probabilities is as follows.

Outcome	Probability
1	$1/6$
2	$1/6$
3	$1/6$
4	$1/6$
5	$1/6$
6	$1/6$

Example 6.3 – Random sampling. A random sample from a known population conforms to the conditions for an ideal game of chance or probability probability experiment. This means that the above definition can be used to determine the probability of selecting a specific individual in a random sample.

A random sample is a means of sampling members of a population so each individual has an equal chance of being selected. That is, no individuals are favoured in the selection procedure and no individuals are discriminated against in terms of the chance of being selected. After the random sample has been obtained, some individuals are selected and others are not, but there is nothing in the structure of the selection process that favours some over others. Researchers often refer to the process of obtaining such a sample as “drawing the sample.”

If a list of all members of a population is available, then such a list can be used to draw a random sample from this population.

If there are a total of N members of the population on the list, then the probability of selecting any specific member on the list is $1/N$.

For example, if there are $N = 40$ students in a class and I randomly select one student from the list of all the students, the probability that any particular student is selected is one in 40, or $1/40 = 0.025$

Events. Across a set of outcomes of a probability experiment, there may be some outcomes with common characteristics. For example, we may be interested in the event of selecting a male or the event of selecting a female when drawing a random sample from a population.

Definition – probability of an event. Suppose a probability experiment has N possible outcomes, with each of these outcomes being equally likely. If an event E has N_E outcomes associated with it, then the probability of event E occurring is N_E/N .

The procedure involved in obtaining the probability of an event is to count the number of equally likely outcomes associated with the event. This is N_E . This count of number of outcomes in the event is divided by the total number of outcomes, N , to obtain the probability that event E occurs.

Example 6.4 – Odd or even when rolling a die. Suppose a single die is rolled. As shown in Example 6.2, each of the six possible outcomes has a probability of one in six of occurring. Suppose the event of an odd number occurring is termed O , and the event of an even number occurring is termed E . Since there are three odd numbers (1, 3, or 5) that can occur when rolling a die, the number of outcomes for which O occurs is $N_O = 3$. Similarly, there are three even numbers (2, 4, or 6) that can occur, so the number of outcomes in event E is $N_E = 3$. From the above definition, the probability of the event of an odd number occurring is $1/2$ or 0.5 .

$$N_O/N = 3/6 = 1/2 = 0.5$$

Similarly, the probability of an even number occurring when a die is rolled is $1/2$ or 0.5 .

$$N_E/N = 3/6 = 1/2 = 0.5$$

Example 6.5 – Males or females in a random sample. Suppose a class of forty students has twenty-five females and fifteen males enrolled in it.

Question. If one student is randomly selected from this class, what is the probability of selecting a female? What is the probability of selecting a male?

Answer. In this case, there are $N = 40$ possible outcomes. Since a student is to be randomly selected from among the whole group, this means that there is equal chance of selecting any student. This selection procedure conforms to the principles associated with an ideal probability experiment, so the above definition of probability associated with an event can be used.

If F is defined as the event of selecting a female, it has $N_F = 25$ possible outcomes associated with it. Using the definition for events, the probability of selecting a female student is $N_F/N = 25/40 = 5/8 = 0.625$.

Similarly, if M is defined as the event of selecting a male, there are $N_M = 15$ males and the probability of selecting a male is $N_M/N = 15/40 = 3/8 = 0.375$.

Some rules and notation for probability

From the above, a few rules concerning probability can be obtained. These are as follows.

- **Between zero and one.** A probability is a number between 0 and 1. This should be apparent from the definitions above. Since each outcome has a probability of $1/N$ of occurring, these are fractions, or numbers, larger than zero but less than one.
- **Impossible event.** If the probability of an outcome or event is zero, it cannot occur. For example, the probability of getting a heart when

rolling a die is zero – this is just not one of the possible events. If the social world, many events are unlikely to occur, but there is often some small chance they could occur. For example, the probability of the Green Party winning the next provincial election is very small, but it might not be zero. A probability of exactly zero is reserved for events that cannot occur.

- **Certain event.** At the other end of the spectrum, an event that is certain to occur is given a probability of one. For example, if I flip a coin and define event HT as that of getting either a head or a tail, then event HT has probability of one. In flipping a coin, I am certain to get either a head or a tail (event HT). In the social world, it is not common to give events a probability of one. That is, there is always some possibility that the event will not occur, so it is not absolutely certain.
- **Sum of probabilities.** The sum of the probabilities of all possible outcomes or events in an ideal game of chance or probability experiment is equal to one. That is, at least one of the outcomes or events must occur. In Example 6.2, rolling a single die, the sum of the six probabilities is 1. That is, the probability of a 1 is $1/6$, the probability of a 2 is $1/6$, and so on, and there are six of these probabilities, so

$$1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1$$

- **Fraction or decimal.** Probabilities can be presented in their fractional form or in decimal form – either is acceptable. In Example 6.5, the probability of selecting a female can be written as either $25/40$ or $5/8$, or as 0.625. When comparing two different probabilities, it is usually easier to compare them when the fractions are converted to decimal form.

Notation. The notation used is that of a mathematical function, in this case a probability function. The notation for the probability of an event E is

$$P(E)$$

If you are unfamiliar with the functional forms, this does not mean P times E , it is just a short form for the phrase “probability of event E ” or simply “ P of E .”

Using this notation, if H indicates a head when flipping a coin and T represents a tail, then the probability of obtaining a head is $1/2$ or 0.5 .

$$P(H) = 1/2 = 0.5$$

and the probability of obtaining a tail is $1/2$ or 0.5 .

$$P(T) = 1/2 = 0.5$$

If a game of chance or probability experiment has N possible outcomes, and if event E has N_E outcomes associated with it, the probability of event E occurring is N_E/N .

$$P(E) = \frac{N_E}{N}$$

The notes in the previous bullets can be restated using the probability notation.

- If E is an event in a probability experiment, then its probability is between zero and one. That is, $P(E)$ is greater than or equal to zero and less than or equal to one.

$$0 \leq P(E) \leq 1$$

- If event E cannot occur, then $P(E) = 0$.
- If event E is certain to occur, then $P(E) = 1$.
- If there are N outcomes for a probability experiment, and these are termed $e_1, e_2, \dots, e_N$, then the sum of the probabilities of these N outcomes is one.

$$\Sigma P(e) = 1$$

Using index notation, if there are n outcomes e_i , where $i = 1, 2, \dots, n$, in a probability experiment, then

$$\Sigma_i P(e_i) = 1$$

The above are a few basic definitions, rules, and procedures for probability. We will return to these rules and further develop them after discussing the other interpretations of probability. The frequency approach (Section B) is examined next, followed by the subjective approach (Section C).

B. Frequency interpretation of probability

Sometimes it is not possible to reason out probabilities using the classical or theoretical approach. For example, if you want to know the probability that it will rain or snow tomorrow, you could sit in an office for an hour thinking about this and you would be no further ahead in estimating this than when you began. Or if you want to know the chance that you will obtain a job that pays well, after you graduate with a bachelor degree, no amount of pure reasoning, thinking, or contemplation will help. What you need to do in each of these situations is obtain some data that is relevant to the issue. For predicting weather, you need to consult weather patterns and maps. To estimate the probabilities associated with getting a good job that pays well, you might look at data about the success or failure of those who have graduated already, and perhaps find a survey of employers that discusses possible job openings in the future.

When applying probability to situations in the world around us, rather than just to ideal games of chance, the frequency interpretation of probability is generally used. This approach constructs probabilities of events based on the number of times that event has occurred in the past. While these probabilities are not perfect guides to the future, they are often preferable to having no estimate of the chance of an event occurring.

Definition – estimate of probability using the frequency approach. If an experiment is conducted n times, and if event E occurs n_E times, then an estimate of the probability of event E is n_E/n .

$$\text{Estimate of } P_E = \frac{n_E}{n}$$

In order to illustrate how this method works, I first use a hypothetical example of flipping a coin.

Example 6.6 – Flipping a coin many times. One way to obtain the probabilities associated with head and tail when flipping a coin is to actually flip the coin many times. Suppose I flip a coin ten times and I get seven heads. The estimated probability of getting a head (H) is then $7/10$ since $n = 10$, $n_H = 7$.

$$\text{Estimate of } P_H = \frac{n_H}{n} = \frac{7}{10} = 0.7$$

But perhaps ten flips is not enough times to obtain a reasonable estimate. So I decide flip a coin one hundred times and suppose I get sixty-three heads. since $n = 100$ and $n_H = 63$, the estimated probability of getting a head (H) now becomes $63/100$.

$$\text{Estimate of } P_H = \frac{n_H}{n} = \frac{63}{100} = 0.63$$

It may be that one hundred flips is still not enough to obtain a good estimate, so I continue the process through one thousand, ten thousand, and one hundred thousand flips of the same coin. Suppose the results are as follows.

Number of flips	Number of heads	Estimate of Probability
10	7	$7/10=0.7$
100	63	$63/100=0.63$
1,000	530	$530/1,000=0.53$
10,000	5,120	$5,120/10,000=0.512$
100,000	50,945	$50,945/100,000=0.509$

If the coin is a fair or honest coin, equally weighted throughout, and if the flipping process is such that outcomes are truly random, then as the number of flips becomes larger and larger, it is likely that the estimated probability gets closer and closer to 0.5000. It may never reach exactly 0.5000, but it may come closer and closer to this value of 0.5000 or one-half. This limiting ratio of one-half can be regarded as the probability of heads.

Definition of probability with frequency approach. If an experiment is conducted n times, and if event E occurs n_E times, then an estimate of the probability of event E is n_E/n . That is,

$$\text{Estimate of } P_E = \frac{n_E}{n}$$

The limiting ratio of n_E/n , as n becomes very large, is defined as the probability of the event E .

Probabilities cited in social science and health research are generally obtained using the frequency approach to probability. If you discuss the probability of a majority government after the next federal election or the likelihood you will get cancer, these are obtained using the frequency approach to probability. In these situations, it is not possible to determine the probabilities on the basis of reason alone, but it is necessary to obtain data about these issues. Such data may be from samples, experiments, or administrative records – in most cases the researcher attempts to obtain a large number of cases in order to determine what the regularities and patterns are. The incidence of a particular outcome across all cases is then interpreted as the probability of the event occurring.

Example 6.7 – Youth in Transition Survey. A Statistics Canada study of young people highlights issues related to the educational and labour market experiences of youth. The study, entitled *At the Crossroads: First Results for the 18 to 20-Year-old Cohort of the Youth in Transition Survey*, contains the following statements.

These findings suggest that the higher the educational level of the parents' education, the more likely their children were to complete high school. (p. 31).

High school graduates were twice as likely as dropouts to have at least one parent who had completed some type of post-secondary education. (p. 31).

Women were somewhat more likely to work a small to moderate number of hours, while men were somewhat more likely to work longer hours. (p. 37).

By using the word “likely,” the authors of the report are placing a frequency interpretation of probability on the results – stating the chances of certain events or combinations of events.

These likely chances, or likelihoods, are probabilities constructed from an examination of a large sample of youth. In the case of this study, “23,592 persons participated in the survey” (p. 70). This was a large sample of Canadian youth from eighteen to twenty

years of age, a sample large enough that the regularities or patterns associated with large sample sizes are stable enough to make conclusions such as the statements from the report. If the sample can be regarded as representative of all youth in Canada, then these statements about likely outcomes can be considered to apply for Canadian youth in general.

Source: The full report is available online at the Statistics Canada web site:

<http://www.statcan.ca/english/freepub/81-591-XIE/free.htm>

The report was written by Jeffrey W. Bowlby and Kathryn McMullen and published by Statistics Canada in 2002.

The frequency approach to probability is extensively used to determine probabilities, especially in social science and health research. In most of these situations dealing with people and their experiences, it is necessary to obtain data from samples, experiments, or administrative sources to determine the incidence of occurrence of events of interest to the researcher. If the number of cases is sufficiently large, and the sample well selected, the fraction of times the event occurs can be interpreted as a probability.

These probabilities derived from the frequency approach can be combined with probabilities obtained from the classical or theoretical approach. Together these different approaches to probability can provide useful insights about human behaviour and social interaction.

C. Subjective or judgment interpretation of probability

The subjective approach to probability involves assigning a number between zero and one to an event. This is sometimes called the judgment approach, in that it involves the judgment or assessment of the situation by the individual assigning the number. Rather than using a systematic means of determining the probability, as was the case in the previous two approaches, this is a subjective judgment made by an individual. It is sometimes referred to as a “degree of belief” that the event will occur.

If you regard an event as very unlikely to occur, then you assign it a small number, close to zero. If you consider the event very likely to occur, you assign a subjective probability of close to one. If you are uncertain whether or no it might occur, you might assign a number such as 0.4 or 0.5 or 0.6 to the event, indicating that it might or might not occur.

Example 6.8 – Hockey lockout. An example of a subjective approach would be to assign a probability to the chance that the National Hockey League lockout will cause the cancellation of the entire season. Since this is not a repeatable event, it cannot be assessed using either the classical or frequency interpretation of probability. But we might want to place a number on the chance that the lockout will continue. If your understanding of the situation is such that you consider it extremely likely that the whole season will be cancelled, then you might assign a number such as 0.9 as its probability. If you are very optimistic that a settlement between owners and players will be reached before the end of 2004, so much of the season can be salvaged, you might assign a probability of 0.1 to the event that all games will be cancelled.

While the subjective approach to probability may be no more than a guess concerning the chances of an event occurring, this can still be useful. If the judgment of the individual providing the estimate is based on reasonable knowledge about, or experience with, the situation, then it can be useful to have a numerical estimate of the chance of occurrence of the event. That is, a well founded judgment can produce a sound estimate of the probability of the event. On the other hand, if an individual has no basis for judgment, or knows little about the situation, then the estimate of probability may

be worthless. As a result, if you use subjective estimates of probability, it is worth determining who made the estimate and how the estimate was obtained.

Subjective estimates of probability can be combined with classical or frequency probabilities to provide some insights into the possibility of occurrence of events. While it can be hazardous to use these subjective probabilities, if thoughtfully constructed by a knowledgeable person, they can be useful.

Recap of approaches to probability

The classical or theoretical approach is the preferred way of determining probabilities – these are well reasoned and well structured probabilities. But they can only be used in particular situations, such as games or chance, ideal probability experiments such as random sampling, and in mathematical formulae. Where possible, these should be used, since these probabilities are precise and can be verified by others.

The frequency approach to probability can also produce precise estimates of probabilities of events under certain conditions. The difficulty with using the frequency approach is that more precise estimates of the chance of events usually require a large and well constructed set of data. That is, in order to provide good estimates of probabilities of events, information must be obtained from many members of a population and the members selected must be somewhat representative of the population. This is often difficult to achieve, but where possible, the incidence of particular events can be interpreted as the chance the event will occur in the future.

If it is not possible to estimate a probability using the classical or frequency approach, then a guess at a number between zero and one can provide a probability. But this judgment or subjective probability is no more accurate than the judgment or guess of the individual making the estimate.

Each of the three types of probability is used in research and in the media. Examples from previous semesters are available on the class web site – look at Problem Set 3 and the associated answers for Winter 2004 or Fall 2003.

Last edited October 21, 2004.