PROBABILITY

Olariu Radu Florin

Regional Department of Defense Resources Management Studies

Abstract:
In this paper are gathered basic concepts used in probability theory, some of the fundamental rules and several uses of them. Most of these concepts are only the prerequisites of more complex models and formulas used in almost every aspect modern life from science to business or even politics. The final goal of these models is to help managers or even individuals in making accurate decisions based on historical facts or data bases that they can access.

Key words: probability, statistics, quantitative, event, random, outcome, impossible, decision.

1. Introduction

Probability theory as a mathematical tool is a very good basis for various human activities related to quantitative analysis from sports (probability of winning a game) to quantum mechanics (probabilistic nature of physical phenomena at atomic scales) or from medicine (drug effectiveness) to politics (probability of winning elections).

Most of the studies are related to risk investment and measure the role of probabilities in making efficient business decisions. We can imagine a large range of situations which need probability algorithms for making the right or the optimum choices.

2. Basic Probability Concepts

To assist making rational and responsible investment decisions, and handle expectations in an environment of risk, an analyst uses the concepts and tools from probability theory. A probability makes reference to the percentage chance that something will occur, from 0 (not possible) to 1 (it is sure to happen), and the scale going from unlikely to probable. Probability concepts help determine risk by quantifying the prediction for unintended and unhelpful results; thus probability concepts are a major focus of numerous studies.

2.1 Random variable

A random variable is defined as any quantity with unknown expected future values. For instance, time cannot be a random variable because we know that a day has 24 hours, a month has 31 days and a year has 365 days and so on. Though, the anticipated rate of return on an investment fund and the expected standard deviation of those returns are random variables. We attempt to predict these random variables taking into consideration past history and on our expectation for the market and interest rates, but we cannot confirm

for sure what values variables will have in the future - all we can come up with are forecasts or anticipations.

2.2 Outcome
Outcome is defined as any potential value that a random variable can take. For estimated rate of return, the range of outcomes reasonably depends on the specific investment or proposition. People who play the lottery have an almost-certain probability of losing all of their investment (-100% return), with an incredibly small chance of becoming a millionaire (+1,000,000% return - or higher!). Consequently for a lottery ticket, there are typically only two extreme outcomes. Mutual funds that invest primarily in stocks will involve a much narrower series of outcomes and a distribution of possibilities around a specific mean expectation. When a particular outcome or a series of outcomes are defined, it is referred to as an event. If our goal for the mutual fund is to produce a minimum 8% return every year on average, and we want to assess the chances that our goal will not be met, our event is defined as average annual returns below 8%. We use probability concepts to ask what the chances are that our event will take place.

2.3 Event
If a list of events is mutually exclusive, it means that only one of them can possibly take place. Exhaustive events refer to the need to incorporate all potential outcomes in the defined events. For return expectations, if we define our two events as annual returns equal to or greater than 8% and annual returns equal to or less than 8%, these two events would not meet the definition of mutually exclusive since a return of exactly 8% falls into both categories. If our defined two events were annual returns less than 8% and annual returns greater than 8%, we've covered all outcomes except for the possibility of an 8% return; thus our events are not exhaustive.

2.4 Probability
Probability has two defining properties:

a) The probability of any event is a number between 0 and 1, or $0 < P(E) < 1$. A P followed by parentheses is the probability of (event E) occurring. Probabilities fall on a scale between 0, or 0%, (impossible) and 1, or 100%, (certain). There is no such thing as a negative probability (less than impossible?) or a probability greater than 1 (more certain than certain?).

b) The sum of all probabilities of all events equals 1, provided the events are both mutually exclusive and exhaustive. If events are not mutually exclusive, the probabilities would add up to a number greater than 1, and if they were not exhaustive, the sum of probabilities would be less than 1. Thus, there is a need to qualify this second property to ensure the events are properly defined (mutually exclusive, exhaustive). On an exam question, if the probabilities in a research study are added to a number besides 1, you might question whether this principle has been met.

These terms refer to the particular approach an analyst has used to define the events and make predictions on probabilities (i.e. the likelihood of each event occurring). How exactly does the analyst arrive at these probabilities? What exactly are the numbers based upon? The approach is empirical, subjective or a priori.

Empirical probabilities are objectively drawn from historical data. If we assembled a return distribution based on the past 20 years of data, and then used that same distribution to make forecasts, we have used an empirical approach. Of course, we know that past

performance does not guarantee future results, so a purely empirical approach has its drawbacks.

Relationships must be stable for empirical probabilities to be accurate and for investments and the economy, relationships change. Thus, subjective probabilities are calculated; these draw upon experience and judgment to make forecasts or modify the probabilities indicated from a purely empirical approach. Of course, *subjective probabilities* are unique to the person making them and depend on his or her talents - the investment world is filled with people making incorrect subjective judgments.

A priori probabilities correspond to probabilities that are objective and derived from interpretation and analysis about a particular case. For example, if we predict that a company is 80% likely to win a bid on a deal (based on an either empirical or subjective approach), and we know this firm has just one business competitor, then we can also make an a priori forecast that there is a 20% probability that the bid will go to the competitor.

Unconditional probability is the direct answer to this question: what is the probability of this one event occurring? In probability notation, the unconditional probability of event A is P(A), which asks, what is the probability of event A? If we believe that a stock is 70% likely to return 15% in the next year, then P(A) = 0.7, which is that event's unconditional probability.

Conditional probability answers this question: what is the probability of this one event occurring, given that another event has already taken place? A conditional probability has the notation P(A | B), which represents the probability of event A, given B. If we believe that a stock is 70% likely to return 15% in the next year, as long as GDP growth is at least 3%, then we have made our prediction conditional on a second event (GDP growth). In other words, event A is the stock will rise 15% in the next year; event B is GDP growth is at least 3%; and our conditional probability is P(A | B) = 0.9.

3. Uses of Probability and Statistics

First come examples of the kinds of practical problems that this knowledge can solve for us. Next this discusses the relationship of probabilities to decisions. Then comes a discussion of the two general types of statistics, descriptive and inferential. Following this is a discussion of the limitations of probability and statistics. And last is a brief history of statistics.

Most important, the chapter describes the types of problems the book will tackle. Because the term "statistic" often scares and confuses people—and indeed, the term has several sorts of meanings—the chapter includes a short section on "Types of Statistics." Descriptive statistics are numbers that summarize the information contained in a group of data. Inferential statistics are procedures to estimate unknown quantities; these procedures infer estimates and conclusions based on whatever descriptive statistics are available. At the foundation of sound decision-making lies the ability to make accurate estimates of the probabilities of future events.

Probabilistic problems confront everyone—from the business person considering plant expansion, to the scientist testing a new wonder drug, to the individual deciding whether to carry an umbrella to work.

PROBABILITY

What kinds of problems can we solve?

These are some examples of the kinds of problems that we can handle with the methods from probability theory:

1. You are a doctor trying to develop a cure for a disease. Currently you are working on a medicine labeled AA. You have data from patients to whom medicine AA was given. You want to judge on the basis of those results whether AA really cures the disease or whether it is no better than a sugar pill.

2. You are the campaign manager for the Libesocial candidate for President of the Romania. You have the results from a recent poll taken in Brasov. You want to know the chance that your candidate would win in Brasov if the election were held today.

3. You are the manager and part owner of a small construction company. You own 20 earthmoving trucks. The chance that any one truck will break down on any given day is about one in ten. You want to know the chance on a particular day – tomorrow - that four or more of them will be out of action.

4. A machine gauged to produce screws 1.000 mm long produces a batch on Tuesday that averaged 1.010 mm. Given the record of screws produced by this machine over the past month, we want to know whether something about the machine has changed, or whether this unusual batch has occurred just by chance.

The core of all these problems is that you want to know the "chance" or "probability that some event will or will not happen or that something is true or false. To put it another way, we want to answer questions about "What is the probability that...?", given the body of information that you have in hand. The question "What is the probability that...?" is usually not the ultimate question that interests us at a given moment. Eventually, a person wants to use the estimated probability to help make a decision concerning some action one might take.

These are the kinds of decisions, related to the questions about probability stated above, that ultimately we would like to make:

1. Should you advise doctors to prescribe medicine AA for patients, or, should you continue to study AA before releasing it for use? A related matter: should you and other research workers feel sufficiently encouraged by the results of medicine AA so that you should continue research in this general direction rather than turning to some other promising line of research? These are just two of the possible decisions that might be influenced by the answer to the question about the probability that medicine AA cures the disease.

2. Should you advise the Libesocial presidential candidate to go to Brasov to campaign? If the poll tells you conclusively that he or she will not win in Brasov, you might decide that it is not worthwhile investing effort to campaign there. Similarly, if the poll tells you conclusively that he or she surely will win in Brasov, you probably would not want to campaign further there. But if the poll is not conclusive in one direction or the other, you might choose to invest the effort to campaign in Brasov. Analysis of the chances of winning in Brasov based on the poll data can help you make this decision sensibly.

3. Should your firm buy more trucks? Clearly the answer to this question is affected by the probability that a given number of your trucks will be out of action on a given day. But of course this estimated probability will be only one part of the decision.

4. Should we adjust the screw-making machine after it produces the batch of screws averaging 1.010 mm? If its performance has not changed, and the unusual batch we observed was just the result of random variability, adjusting the machine could render it more likely to produce off-target screws in the future.

The kinds of questions to which we wish to find probabilistic and statistical answers may be found throughout the social, biological and physical sciences; in business; in politics; in engineering (concerning such spectacular projects as the flight to the moon); and in most other forms of human endeavor.

4. Probabilities and decisions

There are two differences between questions about probabilities and the ultimate decision problems:

1. Decision problems always involve evaluation of the consequences—that is, taking into account the benefits and the costs of the consequences—whereas pure questions about probabilities are estimated without evaluations of the consequences.

2. Decision problems often involve a complex combination of sets of probabilities and consequences, together with their evaluations.

For example: In the case of the contractor's trucks, it is clear that there will be a monetary loss to the contractor if she makes a commitment to have 16 trucks on a job tomorrow and then cannot produce that many trucks. Furthermore, the contractor must take into account the further consequence that there may be a loss of goodwill for the future if she fails to meet her obligations tomorrow—and then again there may not be any such loss; and if there is such loss of goodwill it might be a loss worth $10,000 or $20,000 or $30,000. Here the decision problem involves not only the probability that there will be fewer than 16 trucks tomorrow but also the immediate monetary loss and the subsequent possible losses of goodwill, and the valuation of all these consequences.

5. Conclusion

A working knowledge of the basic ideas of statistics, especially the elements of probability, is unsurpassed in its general value to everyone in a modern society. Statistics and probability help clarify one's thinking and improve one's capacity to deal with practical problems and to understand the world. To be efficient, a social scientist or decision-maker is almost certain to need statistics and probability.

Research in social and physical sciences, statistical testing is not necessary. Where there are large differences between different sorts of circumstances - for example, if a new medicine cures 832 patients out of 1000 and the old medicine cures only 112 patients out of 1000 - we do not need refined statistical tests to tell us whether or not the new medicine really has an effect. And the best research is that which shows large differences, because it is the large effects that matter. If the researcher finds that he must use refined statistical tests to reveal whether there are differences, this sometimes means that the differences do not matter much.

If the raw data are of poor quality, probabilistic and statistical manipulation cannot be very useful. The most refined statistical and probabilistic manipulations are useless if the input data are poor - the result of unrepresentative samples, uncontrolled experiments, inaccurate measurement, and the host of other ways that information gathering can go wrong. Therefore, we should constantly direct our attention to ensuring that the data upon which we base our calculations are the best it is possible to obtain.

References:
[1] www.resample.com

[2] www.intuitor.com/statistics

[3] http://statracket.net/A.%20James%20O'Malley.pdf

[4] http://www.investopedia.com/exam-guide/cfa-level-1/quantitative-methods/probability-subjective-empirical.asp

[5] http://www.math.nyu.edu/faculty/avellane/Lecture1Quant.pdf

[6] http://www.primatech.com/consulting/probability-modeling