IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

USE OF PROBABILITY IN STATISTICS: A STUDY

Dr. Seema Amit Agarwal

Vice Principal

Department of Mathematics / Statistics & Computer Systems

Ghanshyamdas Saraf College of Arts & Commerce

Mumbai, Maharashtra - 400064

ABSTRACT

Probability is commonly used by data scientists to model situations where experiments conducted during similar circumstances, yield different results. Probability allows data scientists to assess the certainty of outcomes of a particular study or experiment. An experiment is a planned study that is executed under controlled conditions. When a result is not already predetermined, the experiment is referred to as a chance experiment. Conducting a coin toss twice is an example of a chance experiment. Today's data scientists need to have an understanding of the foundational concepts of probability theory including key concepts involving probability distribution, statistical significance, hypothesis testing and regression.

Key Word: Probability, Concepts, Interpreting and Rules.

Concepts of Statistics & Probability:

Decisions or predictions are often based on data—numbers in context. These decisions or predictions would be easy if the data always sent a clear message, but the message is often obscured by variability. Statistics provides tools for describing variability in data and for making informed decisions that take it into account.

Data are gathered, displayed, summarized, examined, and interpreted to discover patterns and deviations from patterns. Quantitative data can be described in terms of key characteristics: measures of shape, center, and spread. The shape of a data distribution might be described as symmetric, skewed, flat, or bell shaped, and it might be summarized by a statistic measuring center (such as mean or median) and a statistic measuring spread (such as standard deviation or interquartile range). Different distributions can be compared numerically using these statistics or compared visually using plots. Knowledge of center and spread are not enough to describe a distribution. Which statistics to compare, which plots to use, and what the results of a comparison might mean, depend on the question to be investigated and the real-life actions to be taken.

Randomization has two important uses in drawing statistical conclusions. First, collecting data from a random sample of a population makes it possible to draw valid conclusions about the whole population, taking variability into account. Second, randomly assigning individuals to different treatments allows a fair comparison of the effectiveness of those treatments. A statistically significant outcome is one that is unlikely to be due to chance alone, and this can be evaluated only under the condition of randomness. The conditions under which data are collected are important in drawing conclusions from the data; in critically reviewing uses of statistics in public media and other reports, it is important to consider the study design, how the data were gathered, and the analyses employed as well as the data summaries and the conclusions drawn.

Random processes can be described mathematically by using a probability model; a list or description of the possible outcomes (the sample space), each of which is assigned a probability. In situations such as flipping a coin, rolling a number cube, or drawing a card, it might be reasonable to assume various outcomes are equally likely. In a probability model, sample points represent outcomes and combine to make up events; probabilities of events can be computed by applying the Addition and Multiplication Rules. Interpreting these probabilities relies on an understanding of independence and conditional probability, which can be approached through the analysis of two-way tables.

Connections to Functions and Modeling:

Functions may be used to describe data; if the data suggest a linear relationship, the relationship can be modeled with a regression line, and its strength and direction can be expressed through a correlation coefficient. Probability is often described as “the language of randomness.” The basic idea of probability is that even random outcomes exhibit structure and obey certain rules. In this paper, we will use these rules to build probability models, which are mathematical descriptions of random phenomena. Probability models can be used to answer interesting questions about uncertain real-world systems. Building a probability model involves a few simple steps. First, identify the random variables of interest in the system. A random variable is just a numerical summary of an uncertain outcome. Second, identify the set of possible outcomes for the random variable, which refer to as the sample space. In the example, drawn a ticket from 40 tickets numbered 1 to 40 ,the sample space is the set of numbers from 1 to 40 .

Concept of Probability:

Probability is the measure of the likelihood that an event will occur. Probability is quantified as a number between 0 and 1, where 0 indicates impossibility and 1 indicates certainty. The higher the probability of an event, the more likely it is that the event will occur. A simple example is the tossing of a fair (unbiased) coin. Since the coin is fair, the two outcomes ("heads" and "tails") are both equally probable; the probability of "heads" equals the probability of "tails"; and since no other outcomes are possible, the probability of either "heads" or "tails" is 1/2 (which could also be written as 0.5 or 50%).

These concepts have been given an axiomatic mathematical formalization in probability theory, which is used widely in such areas of study as mathematics, statistics, finance, gambling, science, artificial intelligence/machine learning, computer science, game theory, and philosophy to, for example, draw inferences about the expected frequency of events. Probability theory is also used to describe the underlying mechanics and regularities of complex systems.

Definition and Formula:

Probability is the likelihood of something happening or possibilities of occurrence of an event. When someone tells that the probability of something happening, they are telling that how likely that something is. When people buy lottery tickets, the probability of winning is usually stated, and sometimes, it can be something like 1/10,000,000 (or even worse). This tells that it is not very likely that you will win. The formula for probability tells that how many choices have over the number of possible combinations.

Probability = Possible Choices/Total Number of Options

Calculating Probability

To calculate probability, need to know how many possible options or outcomes there are and how many right combinations are there. Let's calculate the probability of throwing dice to see how it works.

First, we know that a die has a total of 6 possible outcomes. we can roll a 1, 2, 3, 4, 5, or 6. Next, we need to know how many choices we have. Whenever we roll, we will get one of the numbers. We can't roll and get two different numbers with one die. So, our number of choices is 1. Using our formula for probability, we get a probability of 1/6.

Probability = 1/6

Our probability of rolling any of the numbers is 1/6. The probability of rolling a 2 is 1/6, of rolling a 3 is also 1/6, and so on.

Let's try another problem. Let's say we have a grab bag of pomegranate and apple. We want to find out the probability of picking a pomegranate from the bag. One thing we need to know is the number of pomegranates in the bag because that gives us the number of 'correct' choices, which is the number of our possible choices in the top part of the calculation.

We also need to know the total number of fruits in the bag, for this gives us the total number of choices we have, or the total number of options in the bottom part of the calculation. Suppose there are 20 pomegranates and 15 apples in the bag. So, what is our probability of picking a pomegranate? We have 20 pomegranates, one of which we want, and a total of 35 fruits to pick from.

Probability = 20/35

= 4/7

Our probability is 4/7 for picking a pomegranate. If we compare this with our probability of rolling a number on a die, the probability of picking a pomegranate from the grab bag is higher. It is more likely that we will pick a pomegranate than that we will roll a particular number.

The probability of an event that we know for sure will happen is 100% or 1 while the probability of an event that will never happen is 0% or just plain 0.

The probability of these events can be given as a percent or as an odds ratio. Let's pick something a little silly but simple as an example. Let's pretend that we want to wear a shirt to school and we have a red shirt and a pink shirt. The probability of wearing the red shirt is 50% or the odds are 1 out of 2. What is the probability of wearing the pink shirt? It would be the same. The probability of all the events that are possible must add up to 100%.

There are lots of these simple examples that we could use to discuss more about probability. But now we discuss about events important in Life. As we grow up, we need to think about our actions and what the consequences of these actions will be. It's important to know how to use probability when we make decisions about our future. Probability is a whole lot more than just selecting red shirt or blue sweaters or blue socks from the drawer.

A lot of people just wish that they could win the lottery as a way to solve all their financial stresses. In fact, about one-third of the adults in the United States think that winning the lottery is the best way to become financially secure. But we have a better chance of being struck by lightning or being in a plane crash (and there really aren't very many plane crashes) than we have of winning the lottery. The probability of winning the lottery is very close to 0. We need a better plan for our financial future than the lottery!

When we consider probability and our money, we decide how much risk we want to take. The government says that, we will get back your money if the bank itself has financial trouble. If we put money in a bank that is insured by the federal government and we don't exceed the maximum insured amount, the probability of losing our money is 0. But banks usually do not pay very much interest back to us on our savings. We have low risk and low rate of return.

We've seen lots of ads on TV for trading stocks which means buying and selling stock in companies. Before buying a stock, we need to investigate about the company. If the company makes a lot of money with their product and if we own some of their stock we may make more money too, potentially more money than what we could earn at the bank. But if the company loses money, we may lose too. People who work with company finances calculate the probability that a company should make money and are a good company to invest money in.

Another area of our life where probability is important is our health. For example, if we know that people in our family have heart disease and we develop high blood pressure when we are an adult then we know that we have a high probability of also having heart disease. We could be frightened by this high probability or we could live a healthier lifestyle that lowers your blood pressure and in turn lowers the probability of getting heart disease. In this case, we are using our understanding of probability to improve our health.

Use of Probability:

Making Inferences and Justifying Conclusions

-Understand and evaluate random processes underlying statistical experiments

-Make inferences and justify conclusions from sample surveys, experiments and -observational studies

Conditional Probability and the Rules of Probability

-Understand independence and conditional probability and use them to interpret data

-Use the rules of probability to compute probabilities of compound events in a uniform probability model

Using Probability to Make Decisions

-Calculate expected values and use them to solve problems

-Use probability to evaluate outcomes of decisions

Mathematical Practices

-Make sense of problems and persevere in solving them.

-Reason abstractly and quantitatively.

-Construct viable arguments and critique the reasoning of others.

-Model with mathematics.

-Use appropriate tools strategically.

-Attend to precision.

-Look for and make use of structure.

-Look for and express regularity in repeated reasoning.

Interpreting Categorical and Quantitative Data

1.Summarize, represent, and interpret data on a single count or measurement variable

2.Summarize, represent, and interpret data on two categorical and quantitative variables

3.Interpret linear models

Limitations of Probability Functions:

Distributions defined with probability functions differ from those entered with the define assumption command in these ways:

We cannot correlate them.

We cannot view charts or statistics on them.

We cannot extract data from them or include them in reports.

They are not included in sensitivity analyses or charts.

Chances of selecting specific class of samples only

Redundant and monotonous work

generally, probability and statistics help individuals and institutions make better decisions by learning from events in the past and applying this knowledge to the present to affect the future. Probability and statistics are related in that the theories we develop in probability mathematics are compared with statistical findings which can tell us more information about the data. We can also use statistics to estimate the probability of something happening in the future. Probability is a theoretical subject used to analyze the likelihood of events happening in the future. On the other hand, statistics is an applied subject which uses probability theory to analyze data which has been collected

CONCLUSION:

Probability theory seen as logic goes much beyond the so-called frequentists' schools of probability theory, as it is based on more fundamental assumptions (desiderata) and is far more flexible in its applications, as explained in this paper. In fact, probability theory is common sense translated into precise mathematical statements to attain simple and clear results by the process of reduction to calculation. Complexity is reduced to simplicity without neglecting any available information. 'Probability theory as logic' provides a way of drawing consistent conclusions in case of incomplete information that goes well beyond pure inductive or deductive reasoning. Hence, it really stands for a logic in itself.

REFERENCE

Boole, G. 1854. An investigation of the laws of thought, Dover: Macmillan. Reprinted 1958

Carnap, R. 1952. The continuum of inductive methods, University of Chicago Press.

Jaynes, E. T. 1957a,b. Information theory and statistical mechanics. I and II. Physical Review, 106: 20–630. 108, 171–90

Jeffreys, H. 1939. Theory of probability, Oxford University Press. (Later editions: 1948, 1961, 1983.)

Kallenberg, O. (2005) Probabilistic Symmetries and Invariance Principles. Springer -Verlag, New York. 510 pp. ISBN 0-387-25115-4

Kallenberg, O. (2002) Foundations of Modern Probability, 2nd ed. Springer Series in Statistics. 650 pp. ISBN 0-387-95313-2

Olofsson, Peter (2005) Probability, Statistics, and Stochastic Processes, Wiley-Interscience. 504 pp ISBN 0-471-67969-0

Kallenberg, O. (2005) Probabilistic Symmetries and Invariance Principles. Springer -Verlag, New York. 510 pp. ISBN 0-387-25115-4

Kallenberg, O. (2002) Foundations of Modern Probability, 2nd ed. Springer Series in Statistics. 650 pp. ISBN 0-387-95313-2

Olofsson, Peter (2005) Probability, Statistics, and Stochastic Processes, Wiley-Interscience. 504 pp ISBN 0-471-67969-0

Tribus, M. 1996. “The meaning of the word ‘probability’”. In Maximum entropy and Bayesian methods, Edited by: Skilling, J. and Sibisi, S. 143–55. Kluwer Academic.

https://en.m.wikipedia.org/wiki/Probability

https://study.com/academy/lesson/what-is-probability-in-math-definition-lesson-quiz.html

http://googleweblight.com/i?u=http://www.mathworksheetscenter.com/mathtips/calculatingprobability.html&grqid=sxSUnqah&hl=en-IN

http://www.corestandards.org/Math/Content/HSS/introduction/

https://bookdown.org/jgscott/DSGI/probability-models.html