International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

An Analysis of Heuristics and Probability with A Special Reference on Small Sample Fallacy and Representativeness Heuristics

Merlin Mary Thomas[1], Delin Liza Abin[2]

[1]Research Scholar, M.D College, Thrissur
[2]Director, Apple B Academy Private Limited

Abstract

Fallacies and biases are imperative elements while making judgements and decisions. These fallacies often affect the credibility and the final outcome of various research. One such fallacy is the small sample fallacy and the representativeness heuristics. In a small sample fallacy individuals often misunderstand the information derived the small sample, this leads to a cognitive bias. The study focuses on Tossing a coin problem, hospital problem and IQ problem to understand the small sample fallacy and the famous Lind problem is taken to understand the representativeness heuristics. Based on the results, it could be understood that Fallacies of both small sample and heuristics exists even on students with strong statistical background and humans tend to have when it comes to choosing and making decisions, the fallacy that they hold within themselves often tend to mislead them in making day to day decisions

INTRODUCTION:

Heuristics are simple, efficient rules which people often use to form judgements and make decisions. They are mental shortcuts that usually involve focusing on one aspect of a complex problem and ignoring others. Herbert A. Simon proposed that human judgements are based on heuristics, which was further extended by Amos Tversky and Daniel kahneman(1971) which studies how people make real world judgements and the conditions under which these judgements are unreliable. Tversky and Kahneman described a number of heuristics or complex biases like apparent overconfidence, small sample fallacy, regression fallacy, representative and availability fallacies etc that can have an influence in the way people deal with probabilities. By the word fallacy we mean a misconception or a mistaken impression based on unsound arguments. Some fallacies are committed intentionally to manipulate by deception while others are committed unintentionally due to carelessness or ignorance.

As a mathematical foundation of statistics, probability theory is essential to many human activities that involve quantitative analysis of large sets of data. But in real life many people are not familiar with probabilities since it is not much used in everyday life. Uncertainty is an unavoidable aspect of human condition. Many significant choices of our life are based on probability of such uncertain things happening. For example, the result of an election, the outcome of a medical operation, chance of rain etc. By analysing all these situations it is evident that knowingly or unknowingly probability has become an essential element of our life. The problem comes when lay people evaluate the probabilities of uncertain events without actually knowing the theoretical basis which is the underlying cause of various fallacies. While considering probabilities in real life there two rules associated with it- a normative rule and a heuristic

rule. A normative rule relates to an ideal standard or model based on what is considered to be the right way of doing something whereas heuristic rule is a commonsense rule intended to increase the probability of solving some problems that is you try to look at the problem from many angles instead of tackling it head- on.

Through our research we sought to emphasis two major fallacies- small sample fallacy and representative fallacy which is commonly seen in our day to day life. People who commit small sample fallacy can be set to assume that a small random sample should be as reliable as a large random sample which is the heuristic rule. The normative rule is that small samples are not as representative as large sample. People often fail to appreciate the differences in size, relevance of sample size which often forms the foundation for this fallacy.

Tversky and kahneman defined representativeness as the degree to which an event is similar in essential characteristics to its parent population and reflects the salient features of the process my which it is generated. When people rely on representativeness to make judgements, they are likely to judge wrongly because the fact that something is more representative does not actually make it more likely. The representativeness heuristic is simply described as assessing similarity of objects and organising them based around the category prototype.

SMALL SAMPLE FALLACY

In the small sample fallacy for distributions, a small is assumed to be as regular as large sample, or the importance of the difference in regularity is not appreciated. Participants fail to appreciate sample size when judging representativeness.

In a survey conducted among the students of Madras Christian College, seventy samples were collected randomly in order to assess the small sample fallacy. Students with strong statistical background were only considered for the test. The very attempt was to find various ways in which small sample fallacy existed in the minds of people by asking them five different questions.

1. Tossing a coin problem

A question was asked about the likely outcome to occur when a coin was tossed 10 times. The respondents were given two options either choosing one with 9 heads and one tail [HHHHHHHHHT] or heads and tails occurring simultaneously [HTHTHTHTHT]. The response were given in a manner that 72% of the respondents adhered to the first option of a tail occurring after 9 simultaneous heads and 22% sticked on to the second option. Rest 6% said that there is an equal probability of both events happening.

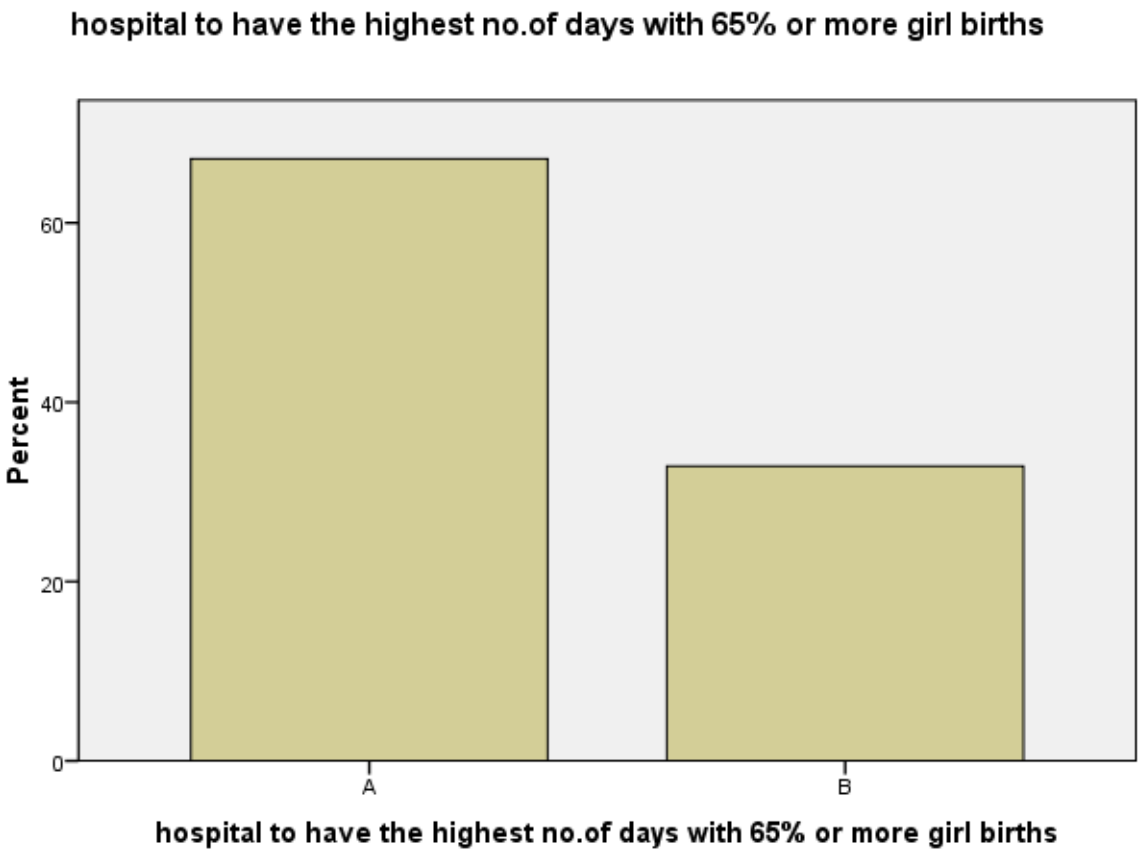
Here students tend to think that after a run of nine heads, there is a probability of tail occurring. According

to the theory, each toss is independent of the previous tosses. This is a special case of gambler's fallacy where people assume that each short segment is as representative as the longer segment.

HOSPITAL PROBLEM

In order to find the problem of small sample fallacy for distribution we relied on one of the examples of Kahneman and Tversky (1972,p.443) where a question was asked regarding the birth of baby girls in hospitals of different size. The respondents were told that two maternity hospitals one being big and other being a small one recorded an average of 40 and 10 child births respectively. They were to decide which hospital that is big hospital or small hospital recorded more days when 65% of the babies born are girls. The three possible choices of answers were the big hospital, the small hospital or both.

Here the correct answer is the smaller hospital because the variability in proportion to their sizes is more for small samples than large samples. 67% of the students gave the option of big hospital and only 32% responded by saying that it was the smaller hospital. However only 1% of students answer about the same. Kahneman and Tversky(1972) present this as an example of the small sample fallacy for distributions. They point out that since the size of the sample does not reflect any property of the parent population, difference in size between the two hospitals should not affect the representativeness or the probability of finding 65% of girls born in each day in the hospital.



IQ PROBLEM

To further ensure the existence of small sample fallacy a question was asked regarding the mean IQ of a population. The question was as follows,

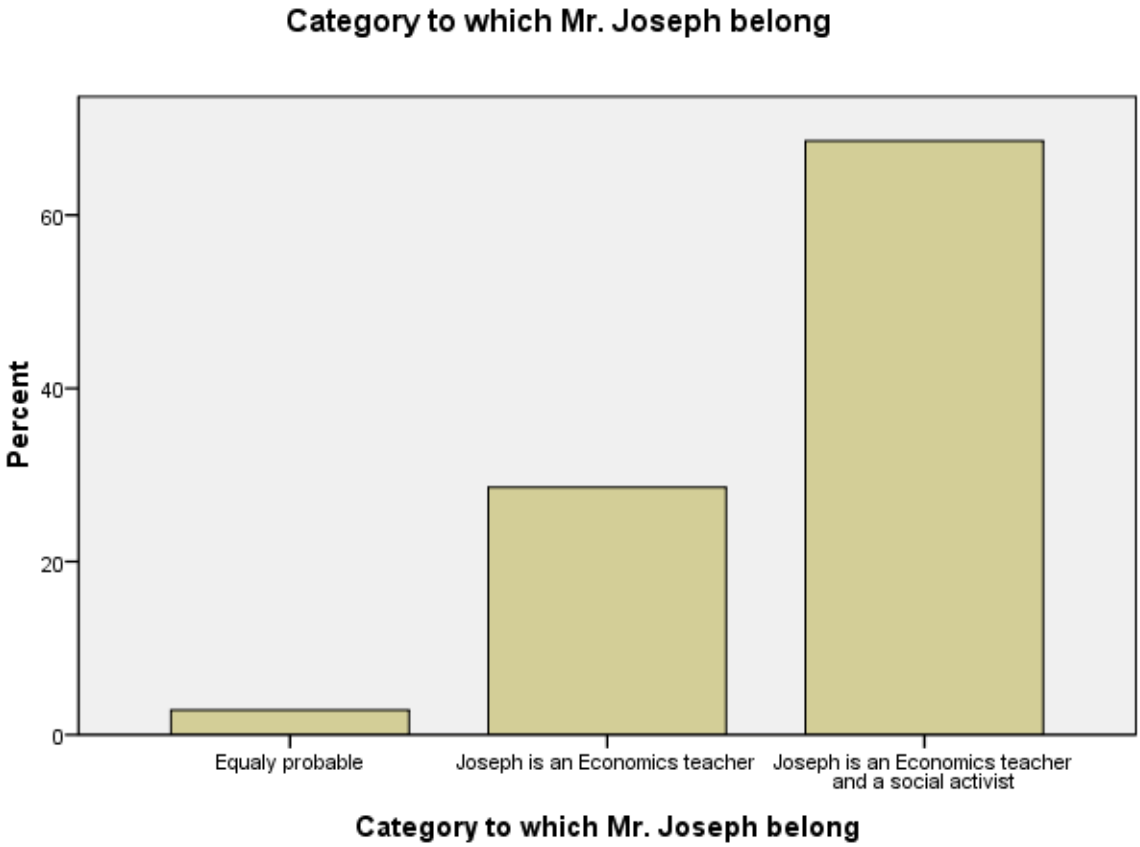
The mean IQ of the population of first year undergraduate students in a college is 100. Random sample of 50 first year undergraduate student was drawn. The first student in the sample scored 150 in the IQ test. What do you expect the mean IQ for the whole sample? Here the correct answer for the question is 101. (150+100*49)/50= 101. Only 5% of the population could give the right answer. And 55% of the population believed that the mean expected IQ is 100. After a score well above 100 they assume that the last 49 scores must average less than 100 in order to compensate for it. This expectation cab be justified only by the belief that a random process is self-correcting. Some familiar processes in nature obey such loss: a deviation from stable equilibrium produces a force that restores the equilibrium [kahneman and tversky 1971]. This is also a kind of gambler's fallacy.

We can safely conclude that small sample fallacy exist even in the mind-sets' of students with statistical background. This fallacy is something even a lay person cannot do away with. This fallacy can be misleading when it comes to taking decisions regarding consumption of a good, business undertakings where we tend to believe that small samples are as representative as large samples. The one who believes in the law of small numbers commits sins against the logic of statistical inference. The belief that small and large samples should be equally reliable and regular(Poultan, 1994, pp78). This is tempered by the belief that small samples ought not to be too regular.

Representativeness

Representativeness reveals a basic tendency to estimate the likelihood of something on the basis of how well it fits a prototype of a familiar category. On the basis of this attribution, people reasoned accordingly and they incorrectly commit conjunction fallacy. In other words people were seduced in their thinking by something known as representativeness. When judging the representativeness of a new stimulus people usually pay attention to the degree of similarity between the stimulus and the standard process. Nilsson, Juslin and Olsson (2008) found this to be influence by the exemplar account of memory (concrete examples of a category are stored in memory) so that new instances were classed as representative if highly similar to a category as well as if frequently encounter.

In order to find out how far the very representativeness is found in a population we took a sample of 70 who had statistical background. The famous Linda problem was taken as a basis for understanding the same and the very first question was based on this where respondents were asked the following question. Joseph is a 35 year old who holds a PG degree in economics. During the student days he was very active in student politics and participated in demonstration against social discriminations. He is also an active member of environmental protection movement. The respondents were asked to choose between two options of Joseph being an economics teacher or Joseph being an economics teacher and a social activist.



Around 69% of the respondents gave the option of Joseph being an economics teacher and a social activist because the description of Joseph has been contrived to convey a person more representative of a social activist than an economics teacher. The description of Joseph matched that of a typical social activist. Joseph's description was representative of a typical social activist therefore on the balance of whole probabilities Joseph must be a social activist. The influence of representativeness was so potent that the participants were drawn to select the incorrect conclusion.

With the base of the same Linda problem another question was asked with slight variation. For the analysis we considered a group of 100 people were 70 are lawyers and 30 are software professionals. The respondents were asked to give the probability that the person being picked at random is software professional. For this part of the question the entire population gave the right answer which was 0.3. But the differences in opinion came for the second part of the question where several attributes about Rachel were given such as Rachel being a married woman and has a child who loves spending free time in attending parties watching movies etc and doesn't show any interest in political and social issues. With these further details the same question was asked to give the chance of Rachel being a software engineer. Here almost 70% deviated from the right answer of 0.3 which shows that increased detailing would mislead people. This is how representativeness would mislead us from taking the appropriate decisions where people get influenced by the increased information.

Heuristics hence provides a good answer quickly but they are not guaranteed to provide the best answer. We tend to forgo careful appraisal of actual probabilities, conditional probabilities and logical relations that actually hold rather takes into account certain intuitions, rules of thumb and biases. The two fallacies

that we have encountered here shows the sensitiveness that humans tend to have when it comes to choosing and making decisions, the fallacy that they hold within themselves often tend to mislead them in making day to day decisions. When it comes to small sample fallacy problem comes when people tend to believe that small samples are as representative as large samples which often is not representative this is what happens when most of the advertisement companies tend to make people believe that a product is worth buying by taking the so called recommendation from a very few doctors and mislead it.

When it comes to representativeness, it entails looking at an event and making a judgement as to how closely it corresponds to other events as found in the general population. They rely on stereotypes. Using representativeness heuristics during problem solving or decision making can give rise to several fallacies. These fallacies can lead to poor decision making or problem solving for the person that falls victim to this way of thinking.

In complex situations, people often rely on certain shortcuts to reach conclusion or decision. However, if the shortcuts don't represent the situation adequately, it can lead to misrepresentation of probability, statistics and nature of chance situations.

REFERENCE

1. Poulton,E.C, (1994). Behavioural decision theory: A new approach. Cambridge university press, 5, 78-104
2. Quinlan Philip& Dyson Ben, (2008). Cognitive Psychology, 15, 557-559
3. Tversky, A & kahneman, D,(1971). Belief in the law of small numbers. Psychological bulletin, 76, 105-110
4. Tversky, A & Kahneman, D, (1972). Subjective probability: A judgement of representativeness. Cognitive Psychology, 3, 430-454
5. Tversky, A & Kahneman, D (1973). Availiability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232