BIAS

#3 Information, part 1

UNDERSTANDING AND OVERCOMING BIAS

I have found with my work and research about bias that I now notice them continually in not only my consulting, but just in everyday life. For example, last year I was teaching R&A's Mitigating Bias course for a client and coincidently while I was at breakfast, I picked up the paper and ran across an article on vaccines and some people's reticence about immunizing themselves or their children.

Interestingly, much of the article talked about the various cognitive biases we have, and how they impact our decisions. There were several that were mentioned, but one I want to talk about in this blog is how people struggle with data, statistics, probability, and numbers. For example, the article noted how "people are flummoxed by numerical risk, paying more attention to numerators, such as '16 adverse events' than we do to denominators such as 'per million vaccine doses'" (New York Times International, September 26, 2019).

***DATA, DATA EVERYWHERE***

In today's world, we are inundated with so much information and data, that it often can be difficult to deal with. As we talked about in my previous blog article, we use heuristics to try and make sense of the world around us. Remember, these heuristics are meant to be simplified and quick shortcuts in our decision making and judgment. Interestingly, there is a branch of statistics known as 'intuitive statistics' which refers to our ability to make generalisations and predictions using our intuition and experience, reinforced by our heuristics.

Unfortunately, as noted before, this can lead to cognitive biases and systematic errors in judgements and decisions. One such bias that affects our judgments of data is the *information* bias which reflects our tendency to have a distorted perception of information and its significance. Put simply, statistics and probability are just not intuitive!

In one respect, we are not really interested in the numbers; rather we are interested in the outcome. For example, a 60% probability that it will rain tomorrow does not do me much good. What I am interested in is whether I will get wet and if I should bring an umbrella! Hence, we will make judgments and decisions seeking certainty (for example, an observation I've made to support this view is that people will often convert probabilities of occurrence that are <40% to 0% and ones that are >60% to 100% for just this reason).

***MONTY HALL PROBLEM***

Let's take a look at a famous example popularized in the "Ask Marilyn" question-and-answer column of the Parade magazine (see "Power of Logical Thinking" by Marilyn vos Savant, St. Martin's Press, 1996), and named after a famous game show in the United States.

*Suppose you are on a game show, and the game show host gives you the choice of three doors. Behind one door is $10,000, and behind the other two doors are goats; and you are trying to win the money. You pick a door, say number 1. The host, who knows what is behind each door, purposefully opens another door that only has a goat behind it – say door #3. He then asks you "Do you wish to keep your original door #1 pick, or would you like to switch to door #2?" Is it to your advantage to switch doors, not to your advantage, or it does not matter?*

If you have not seen this problem before, think of your answer. Most people say it does not matter as there is a 50-50 chance that the money is behind any one of the two remaining doors. This is not the correct answer, and actually you have a 66.67% chance of winning if you switch doors! I will leave it to you to research why, but suffice to say that Marilyn's correct response (note, at the time she had one of the highest IQ's in the world) caused an avalanche of correspondence, mostly from people who said her solution was wrong (including 65% who had a university degree, and 1000 with a PhD!) – one sample letter stated: "May I suggest that you obtain and refer to a standard textbook on probability before you try to answer a question of this type again?" (University of Florida, PhD)!

This apparently simple, but surprisingly complex, example illustrates how our intuition lets us down.

***THIS INFORMATION IS SURELY REPRESENTATIVE***

A shortcut that we often use to make judgments and simplify decision making about information is the *representative* heuristic, where we make judgments on how similar something is perceived to be to something else, often ignoring statistical evidence.

This heuristic leads us to systematically poorly predict the likelihood of an event and quantify uncertainty. We have a poor sense of what constitutes a random sequence, tend to underestimate the relevance of sample sizes, ignore base rates and dependencies (or not) between variables, are insensitive to prior probabilities and outcomes, and subjectively weight probabilities. Our judgments of data and probability are often distorted by our other cognitive bias. For example, the *availability* bias can incorrectly lead you to believe that the frequency that you have encountered something (particularly if it is an extreme or dynamic event) will imply that it is more likely to occur, which can be further reinforced *anchoring* and *confirmation* biases.

Let's take a look at a few of the *representative heuristics'* pitfalls.

***RANDOMNESS***

One of my favorites is how people underestimate the frequency and impact of randomness! Apple actually reprogrammed their iTunes shuffle feature to make it appear more random!

I remember when I was in university and taking a statistics class the professor would divide the class into two; one half he would have the class flip a coin 200 times and record their answers and the other half he told them to write down their own sequence of 200 flips. Without fail he could always tell which one of the groups actually flipped the coin and which ones were human generated – he just looked for a sequence of 5 or more. The human generated group would resist generating a long sequence because that is not their expectation of randomness. Perhaps to your surprise, the probability when flipping a fair coin 200 times of having at least one string of 5 or longer heads or tails is 99.9%!

***BASE RATE***

Our tendency is to focus on specific information and ignore what is generally true (i.e., *base rate*). In other words, we tend to ignore the probability of a specific phenomenon occurring, and rather focus on our specific observations of the event (which can be reinforced by the *availability* bias). For example, when we see a product advertisement of "50% extra free", our intuition thinks that half of the item is free; whereas when taking the base rate into account, it only implies that 1/3 of the item is free.

Bayes Theorem was applied by the US Government in 2009 to advise women in their forties with no relevant history not to have a routine, annual mammogram. The public reaction was immediate and enraged! Why did the US Government make this recommendation, and what were the public not considering? Answer next blog!

***SAMPLE SIZE***

People tend to believe that a sample population can be adequately represented with a selective or small number of data points and hence, place too much confidence in a small sample being representative of the general population**.** By its nature, a small sample is more likely to deviate markedly from the average of the population from which it is drawn; and although we know a large sample is better, we do not intuitively understand the difference between a small and large sample size and tend to equate them and our conclusions equally. Outcomes from a small sample can often look unrepresentative (e.g., flip 5 heads in a row). As noted above, that is not a biased outcome.

Interestingly, there are two aspects to the *sample size* heuristic – size and bias. We often just think of the former (i.e., how big of a sample we have); yet often a sample can be biased, which is a significant issue for pollsters. In our industry, it could manifest itself by using samples for our reservoir parameters only from one successful, adjacent field.

***UNTIL NEXT TIME***

On my next blog article, I will further explore the *information* bias and *representative* heuristic by interviewing two experts in our field. Peter Carragher (Rose & Associates) and Graeme Keith (Stochastic ApS) will give us some more insights on why we struggle with statistics and probability, and strategies on how to overcome.

Stay safe and healthy.

*Marc Bond*
*Rose & Associates*