Entry: Base Rate Fallacy
Word Count: 991
Author: Nicholas Epley

Definition of term

Imagine that you meet Tom one evening at a party. He is somewhat shy and reserved, is very analytical, and enjoys reading science fiction novels. What is the likelihood that Tom works as a computer scientist? The answer depends on both the knowledge you have about Tom, but also on the number of computer scientists that exist in the population. Tom fits the stereotype of a computer scientist, but there are relatively few computer scientists in the general population compared to all other occupations. The knowledge you have about Tom is often called individuating or case-based information, whereas knowledge about the number of computer scientists in the general population is often called distributional or base rate information. When presented with both pieces of information—be it when judging the risk of contracting a disease, when judging the likelihood of a defendant's guilt, or when predicting the likelihood of future events—people often base their judgments too heavily on case-based or individuating information and underutilize or completely ignore distributional or base-rate evidence. Underutilizing or ignoring base-rate evidence in intuitive judgments and decision-making is known as the *base rate fallacy*.

Background

The classic scientific demonstration of the base rate fallacy comes from an experiment, performed by psychologists Amos Tversky & Daniel Kahneman, in which participants received a description of 5 individuals apparently selected at random from a pool of descriptions that contained 70 lawyers and 30 engineers, or vice versa. Participants were asked to predict whether each of the 5 individuals was a lawyer or an engineer. The compelling result was that participants' predictions completely ignored the composition of the pool (i.e., the base rates) from which the descriptions were drawn. Instead, participants seemed to base their predictions of each person's occupation on the extent to which the description resembled, or was similar to, the prototypical lawyer or engineer. Relying on this *representativeness heuristic* led participants to completely regard the base rates that should also have been incorporated into their predictions.

Results like these have been replicated in a wide variety of contexts since this initial demonstration. Underutilizing population base rates has been used, for instance, to explain why people are overly concerned about extremely rare events (such as dying in a terrorist attack or contracting a rare disease), why people pay for insurance they do not need, and why doctors misdiagnose their patients. However, broad conclusions about the general existence and robustness of the base rate fallacy in daily life have become quite controversial for two reasons. First, experimental results often show that people do indeed utilize base rates when making some predictions in at least some contexts. Empirical research simply does not support the claim that people completely ignore base rate evidence when making judgments and decisions. Second, statisticians have pointed

out the difficulty in determining exactly how much people should incorporate base rates into their judgments in daily life. It is therefore difficult, in some contexts, to argue that people *should* incorporate base rates into their judgments and decisions that they naturally ignore or apparently underutilize.

Evidence

Empirical evidence suggests that base rates are sometimes completely ignored and at other times appropriately utilized. The key issue for social psychologists, then, is to understand when the base rate fallacy is likely to emerge and when it is not. At least four major factors are known to moderate people's use of base rates in judgments and decisions.

First, people are more likely to utilize base rates when making multiple judgments with different base rates than when making a single judgment. People would be more sensitive to the base rates of diseases in the population, for instance, when judging the likelihood that they will contract AIDS, Influenza, and Malaria then when judging the likelihood that they will contract any of these diseases in isolation.

Second, people are more likely to use base rates when they have little individuating or case-specific information to use in its place. People are more likely to utilize base rates, for instance, when predicting the behavior of a randomly selected person than when predicting their own behavior, in large part because little individuating or case-based information is available for the "random person" but a good deal of individuating information is present when predicting one's own behavior.

Third, people are more likely to utilize base rates when they are perceived to be valid and reliable. Base rate information about elderly adults, for instance, is more likely to be utilized when making judgments about elderly adults than when making judgments about young adults. Base rates tend to be ignored when they are perceived to be invalid and unreliable.

Finally, people are more likely to use base rates when they are presented as frequencies of occurence (e.g., 1 out of 100) than when they are presented as single-case probabilities or occurence (e.g. 1%). People would be more sensitive to the actual population base rates, for instance, when predicting how many commercial airplane flights out of 1,000 will crash due to mechanical malfunctions than when predicting the likelihood (from 0% to 100%) that any single airplane flight will crash due to mechanical malfunctions.

Importance

Both trivial and important decisions are often based on the perceived likelihood of events. People avoid flying if they believe the likelihood of a crash is high, marry a dating partner if they believe the likelihood of divorce is low, and start new businesses depending on the perceived likelihood of success. Nearly all likelihood judgments require the integration of case-based or individuating information and distributional or base rate evidence. Understanding when people are likely to utilize these base rates appropriately versus inappropriately provides insight into when people are likely to make good versus bad decisions, and understanding why people might sometimes commit the base rate fallacy provides insight for how to improve everyday decision making.

Additional Readings:

Koehler, J.J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*, 1-53.