



The Language of Statistics (and What's Lost in Translation)

Galit Shmueli^{1*}

¹ Institute of Service Science, National Tsing Hua University, Hsinchu, Taiwan
galit.shmueli@iss.nthu.edu.tw

Abstract:

The field of statistics uses “statistical language” for describing phenomena in the observable world. This language includes statistical notation for parameters, sample statistics, random variables and their realizations, as well as functions of these objects such as densities and conditional probabilities. While statistical language has proven powerful for both theory development and practical applications, some phenomena of interest to researchers and practitioners in business, industry, and otherwise, are not easily conveyed using statistical language, thereby creating statistical blind spots. Identifying such “unknown unknowns” can expand statistical research and applications in new domains. I will describe three blind spots encountered when collaborating with management scientists, behavioral scientists, and other research communities. One blind spot relates to *concepts* and *constructs* — abstractions of interest in many scientific fields. The second relates to analysis *goals* and *utility*, which are key to statistics in practice. The third blind spot concerns *causal interventions* and *feedback loops*, especially in combination with prediction.

Keywords:

statistical notation; concepts; analysis goal; feedback loop; causal intervention