Title: Communicating Uncertainty Through Visualization

Submitters:

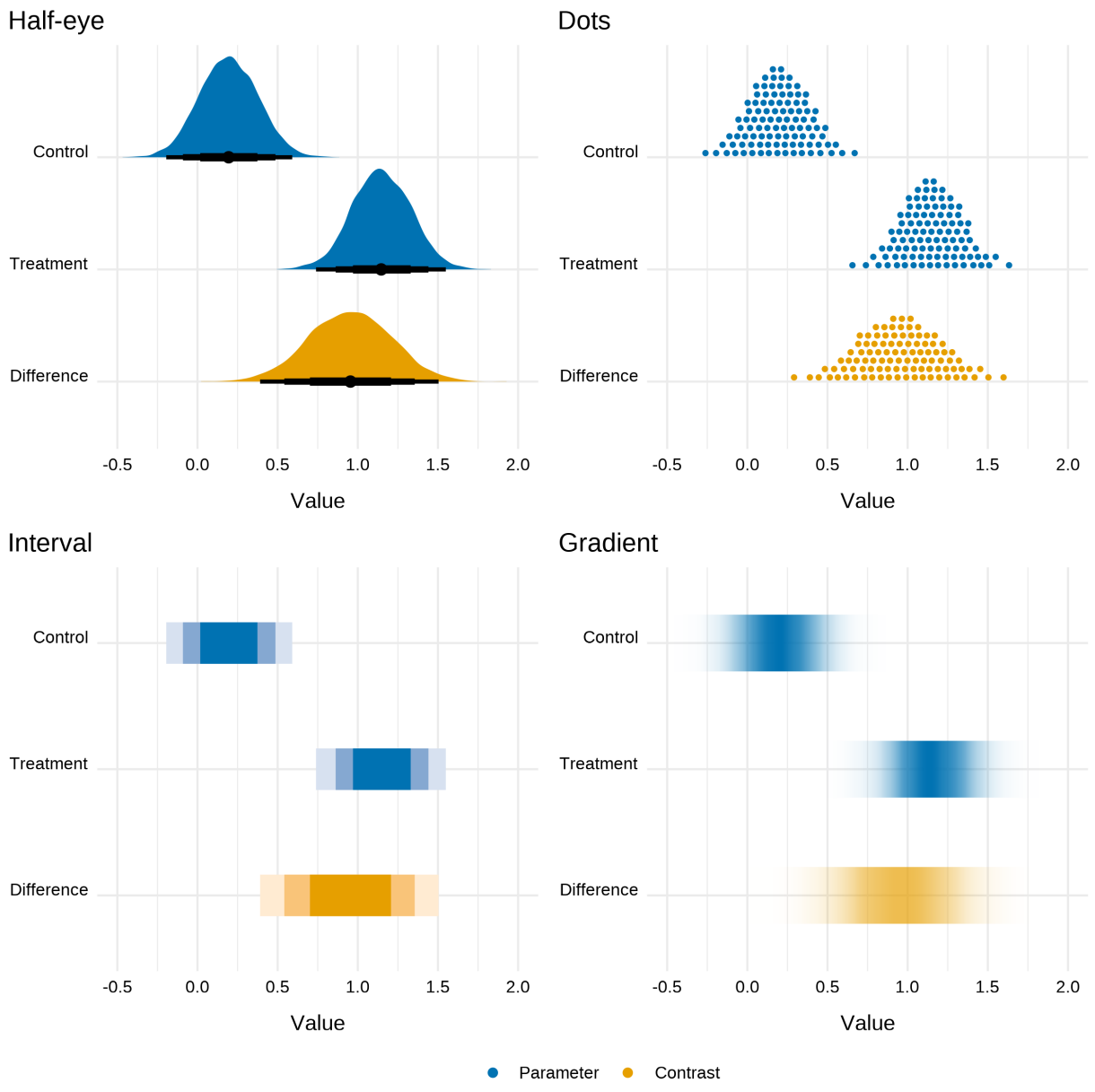
W. Jake Thompson, Ph.D.
Accessible Teaching, Learning, and Assessment Systems (ATLAS); University of Kansas

Corresponding author:

W. Jake Thompson, Ph.D.
ATLAS, University of Kansas
1122 West Campus Road
Joseph R. Pearson Hall, Room 437
Lawrence, KS 66045
wjakethompson@gmail.com

Title: Communicating Uncertainty Through Visualization

Caption: Four different methods for visualizing uncertainty in parameter estimates.



Description:

Accurately and effectively communicating uncertainty is a critical aspect scientific research. Often, results of scientific studies are reported as an average effect or model parameter with an associated standard error or confidence interval. However, this type of reporting often obscures important aspects of the data, and confidence intervals written in text are easy to skim over. Communicating uncertainty through visualization allows consumers of scientific research to see a more complete picture of the data or model, and can lead to more effective decision making (Padilla et al., 2023). This graphic uses simulated data to illustrate four different methods for visualizing the uncertainty in group-level effects and the difference between the effects. These uncertainty distributions are straightforward to create using the {ggplot2} and {ggdist} packages in R (Kay, 2022; Wickham et al., 2022).

Half-eye plots use density plots with interval lines to visualize the complete distribution of plausible parameter values, while the interval lines allow readers to more easily summarize the high density areas. Dot plots (also called quantile dot plots) are similar to density plots, but summarize the distribution into a given number of dots or quantiles (100 in this graphic). This method of representing distributions as frequency counts has been shown to be easier for laypeople to understand (Kay et al., 2016). Interval plots summarizes ranges of plausible values (e.g., 67%, 89%, and 97% in this graphic). The wider intervals (i.e., more uncertainty) are shaded with a lighter color, where as the narrower intervals are darker. Finally, gradient plots are similar to interval plots, but rather than having sharp boundaries at arbitrary thresholds, the gradient plot is continuous. Transparency is used to "fade out" the low density regions, whereas the high density regions are opaque.

References

Kay, M. (2022). *ggdist: Visualizations of distributions and uncertainty*. R package version 3.2.0.https://mjskay.github.io/ggdist

Kay, M., Kola, T., Hullman, J. R., & Munson, S. A. (2016, May 7–12). *When(ish) is my bus? User-centered visualizations of uncertainty in everyday, mobile predictive systems* [Paper presentation]. CHI '16: CHI Conference on Human Factors in Computing Systems (pp. 5092–5103), San Jose, CA. https://doi.org/10.1145/2858036.2858558

Padilla, L., Kay, M., & Hullman, J. (2023). Uncertainty visualization. In N. Balakrishnan, T. Colton, B. Everitt, W.Piegorsch, F. Ruggeri, & J. L. Teugels (Eds.), *Wiley StatsRef: Statistics reference online*. Wiley. https://doi.org/10.1002/9781118445112.stat08296

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (2022). *ggplot2: Create elegant data visualisations using the grammar of graphics*. R package version 3.4.0. https://ggplot2.tidyverse.org