Profits of Prejudiced Algorithms: README

David J. Jin[†]

Remark. This data file contains 34 files: 4 R scripts and 30 simulated data sets.

- "sims_all.R": R script that conducts simulations needed to generate Figure 5
  - "maintestdf_1," "maintestdf_2," and "maintestdf_3" are synthetic "testing data sets" corresponding to normal, exponential, and gamma distributions, respectively, for the data-generating processes for applicant signals.
- "sims_norm.R": R script that conducts simulations needed to generate Figure C1 (Appendix)
  - This script corresponds to nine synthetic data sets "normtestdf_x_y" where $x, y \in \{1, 2, 3\}$. The coordinate x is in bijection with three levels of $\mathbb{P}(Q = 1) = \pi \in \{0.4, 0.5, 0.6\}$. The coordinate y is in bijection with three values of the mean of signals of the qualified population $\{1, 2, 3\}$.
- "sims_exp.R": R script that conducts simulations needed to generate Figure C2 (Appendix)
  - This script corresponds to nine synthetic data sets "exptestdf_x_y" where $x, y \in \{1, 2, 3\}$. The coordinate x is in bijection with three levels of $\mathbb{P}(Q = 1) = \pi \in \{0.4, 0.5, 0.6\}$. The coordinate y is in bijection with three values of the inverse mean of signals of the qualified population $\{2, 3, 4\}$.
- "sims_gamma.R": R script that conducts simulations needed to generate Figure C3 (Appendix)
  - This script corresponds to nine synthetic data sets "gammatestdf_x_y" where $x, y \in \{1, 2, 3\}$. The coordinate x is in bijection with three levels of $\mathbb{P}(Q = 1) = \pi \in \{0.4, 0.5, 0.6\}$. The coordinate y is in bijection with three values of the mean of signals of the qualified population $\{2, 3, 4\}$.

[†]Yale University; d.jin@yale.edu.