Explanatory Models for Understanding Differential Item Functioning

Skorupski, W. P., Loughran, J. T., & Brussow, J. A.

National Council on Measurement in Education, Chicago, IL, April 2015

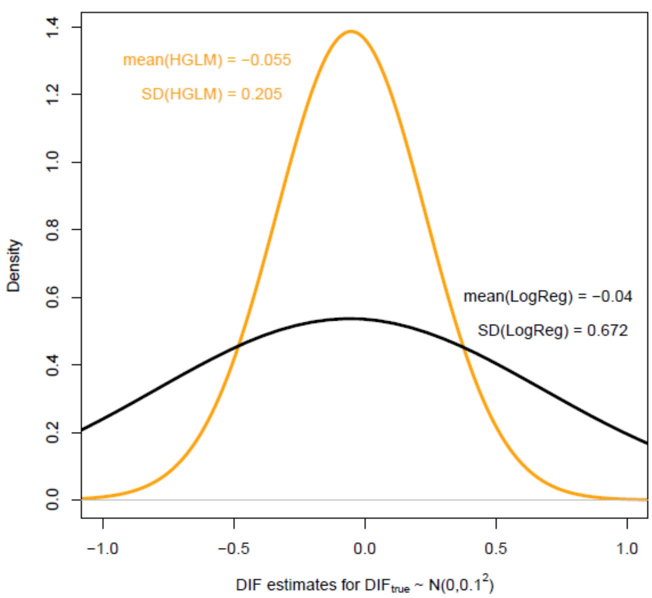
Purpose

- Evaluate use of item-level features as explanatory variables for understanding DIF
- Understanding item-level features' impact could change the way we write items to avoid DIF

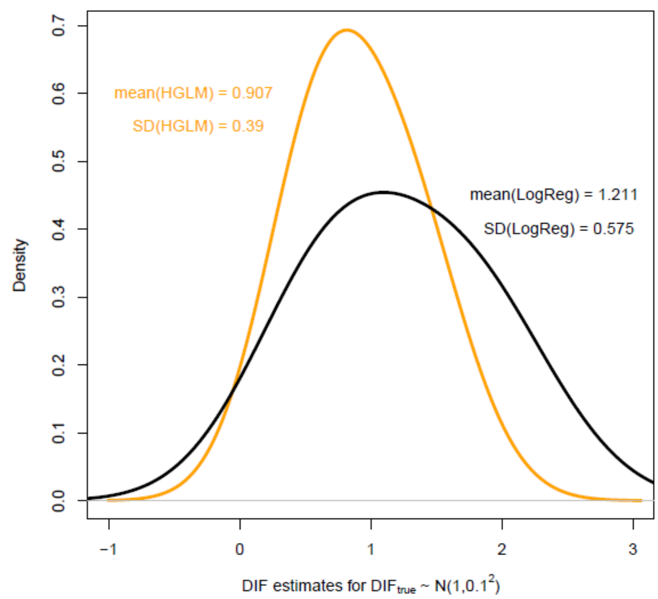
Methods

- Simulated data from 3-PL IRT model
  - 1,000 examinees taking 40-item test; Focal and reference groups: N=500
  - Reference group θ ~N(0, 1); Focal group θ ~N(-0.5, 1)
  - Difficulty: b ~N(0, 1); Discrimination: a ~U(0.5, 3.5); Lower asymptote: c ~U(0.1, 0.3)
  - 2 items with DIF: D ~N(1, 0.1^2); 38 items w/ random, negligible DIF: D ~N(0, 0.1^2)
- Comparing 100 simulated datasets using:
  - Logistic regression followed by ordinary least squares regression (two-stage)
  - Fully Bayesian Hierarchical Generalized Linear Model (simultaneous)

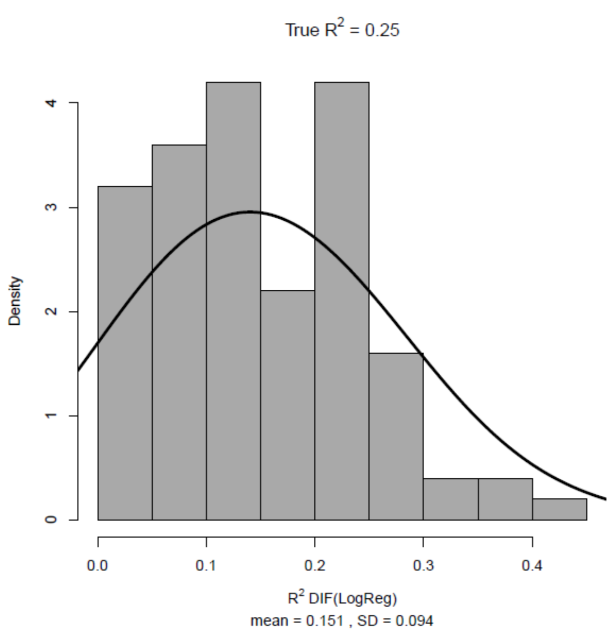
Density plots of DIF estimates for items simulated to have negligible DIF.



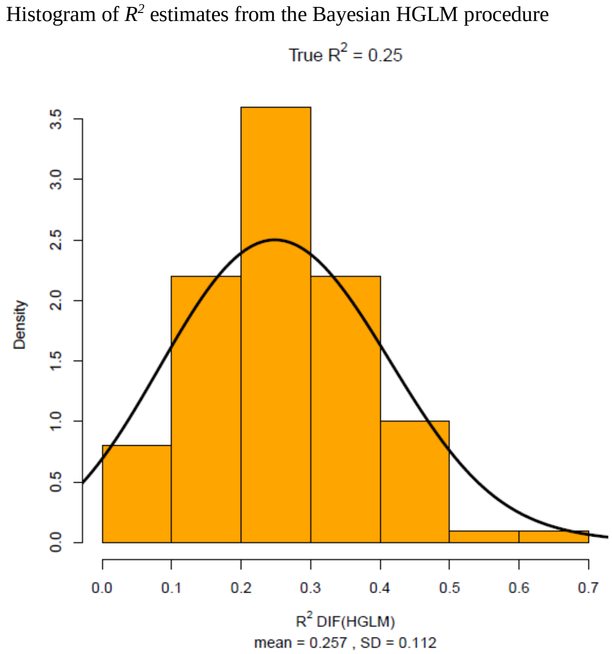
Density plots of DIF estimates for items simulated to have meaningful DIF.



Histogram of R^2 estimates from the linear regression of DIF estimates (two-stage)



Histogram of R^2 estimates from the Bayesian HGLM procedure



- Bayesian HGLM is a reliable procedure for evaluating & interpreting DIF with respect to interpretable item features
- Bayesian HGLM was more precise and accurate in estimating DIF & relationship (i.e., R^2) between DIF & item features