binGroup: A Package for Group Testing

Christopher R. Bilder[1⋆], Boan Zhang[1], Frank Schaarschmidt[2], Joshua M. Tebbs[3]

1. Department of Statistics, University of Nebraska-Lincoln, Lincoln, NE
2. Institute of Biostatistics, Leibniz University Hannover, Germany
3. Department of Statistics, University of South Carolina, Columbia, SC

⋆Contact author: chris@chrisbilder.com

Keywords: Binary response, Generalized linear model, Group testing, Latent response, Pooled testing

When the prevalence of a disease or of some other binary characteristic is small, group testing is frequently used to estimate the prevalence and/or to identify individuals as positive or negative. Examples of its use include human infectious disease detection, veterinary screening, drug discovery, and insect vector pathogen transmission rate estimation (Pilcher et al., 2005; Peck, 2006; Remlinger et al., 2006; Tebbs and Bilder, 2004). We have developed the **binGroup** package as the first package designed to address the estimation problem in group testing. We present functions to estimate an overall prevalence for a homogeneous population. Also, for this setting, we have functions to aid in the very important choice of the group size. When individuals come from a heterogeneous population, our group testing regression functions can be used to estimate an individual probability of disease positivity by using the group observations only. We illustrate our functions with data from a multiple vector transfer design experiment and a human infectious disease prevalence study.

References

C. Peck. (2006). Going after BVD. *Beef*, 42, 34–44.

C. Pilcher, S. Fiscus, T. Nguyen, E. Foust, L. Wolf, D. Williams, R. Ashby, J. O'Dowd, J. McPherson, B. Stalzer, L. Hightow, W. Miller, J. Eron, M. Cohen, and P. Leone. (2005). Detection of acute infections during HIV testing in North Carolina. *New England Journal of Medicine*, 352, 1873–1883.

K. Remlinger, J. Hughes-Oliver, S. Young, and R. Lam. (2006). Statistical design of pools using optimal coverage and minimal collision. *Technometrics*, 48, 133–143.

J. Tebbs and C. Bilder. (2004). Confidence interval procedures for the probability of disease transmission in multiple-vector-transfer designs. *Journal of Agricultural, Biological, and Environmental Statistics*, 9, 75–90.