Biometrika (2018), **yyy**, y, *pp*. 1–2
Printed in Great Britain

Discussion of Sesia et al. 2018 and the Knockoff Framework

BY JONATHAN D. ROSENBLATT
Dept. of Industrial Engineering and Management,
Ben Gurion University of the Negev, Israel.
johnros@bgu.ac.il

AND JELLE J. GOEMAN
TOOD: Jelle
J.J.Goeman@lumc.nl

1. ON THE PROBLEM

The authors of Sesia et al. (2018) set out to design a procedure for variable selection with provable statistical guarantees. The *knockoff* algorithm proposed by Sesia et al. (2018), provably controls the FDR of conditionally independent variables. Denoting with x and y the predictor and outcome variables, respectively. The *false detection proportion*, a.k.a. the *false selection proportion*, is defined as V/R where R is the number of variables selected, and V is the number of such variables where $y|x_{-j}$ is independent of x_j. The knockoff algorithm of Sesia et al. (2018) provably control the $FDR := \mathbb{E}[FDP]$, at some user selected magnitude.

Crucially for our comments: (1) The FDR is an expectation with respect to variability in x and y, i.e., a *random design* guarantee. (2) The procedure is *model free*, or *non-parametric* in that nothing is assumed on the parametric form of $y|x$. (3) The proofs assume full knowledge of F_x, i.e., the joint distribution of predictors, marginalized over y. (4) The method aims at good variable selection, not prediction.

2. ON THE METHOD

The fundamental idea of the method is to generate variables that have all the properties of the original x_j, only that they are conditionally uncorrelated to y. These are termed *knockoff* variables. The method then proceeds to compute a test statistic that captures the difference in the strength of the dependence of the knockoff, and the original variable. This statistic is then compared to it resampling distribution: the distribution over resampled knockoffs.

2·1. *Other Variable Selection Methods*

2·2. *Other Knockoffs*

2·3. *Permutation Testing and Symmetries*

3. FUTURE RESEARCH

4. ON THE HYPOTHESES

5. ON THE PROBLEM SETUP

ACKNOWLEDGEMENT

The authors thank Prof. Yaakov Ritov, Dr. Aldo Solari, Dr. Livio Finos, and ... for fruitful discussions leading to this manuscript.

REFERENCES

SESIA, M., SABATTI, C. & CANDES, E. J. (2018). Gene hunting with hidden markov model knockoffs. *Biometrika* 00000.

[*Received* y yyy yyyy. *Editorial decision on* y yyy yyyy]