Supplement to 'Testing the significance of interactions in genetic studies using interaction information and resampling technique'

Paweł Teisseyre *[1], Jan Mielniczuk [1, 2], and Michał J. Dąbrowski[1]

[1]Institute of Computer Science, Polish Academy of Sciences, Poland
[2]Warsaw University of Technology, Faculty of Mathematics and Information Science, Poland

1 Example of positive and negative interaction information.

2 Simulation models for power comparison

Simulation model M1:

$$s(y) = \begin{cases} p, & \text{if } y = -1, \\ 0, & \text{if } y = 0, \\ p, & \text{if } y = 1. \end{cases}$$

Simulation model M2:

$$s(y) = \begin{cases} 0, & \text{if } y = -1, \\ p, & \text{if } y = 0, \\ p, & \text{if } y = 1. \end{cases}$$

Simulation model M3:

$$s(y) = \begin{cases} 0, & \text{if } y = -1, \\ p, & \text{if } y = 0, \\ 0, & \text{if } y = 1. \end{cases}$$

Simulation model M4:

$$s(y) = \begin{cases} 0, & \text{if } y = -1, \\ 0, & \text{if } y = 0, \\ p, & \text{if } y = 1. \end{cases}$$

3 Real data analysis

*Correspondence to: Paweł Teisseyre, Institute of Computer Science, Polish Academy of Sciences, 5, Jana Kazimierza, 01-248 Warsaw, Poland, e-mail: teisseyrep@ipipan.waw.pl

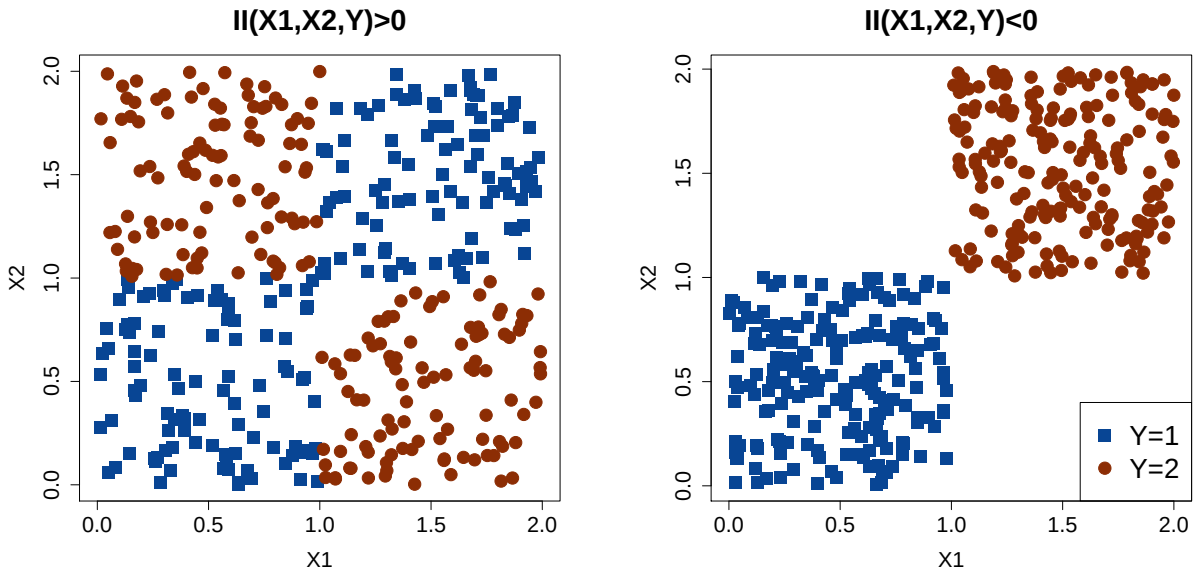


Figure 1: Examples of positive interaction information (a) and negative interaction information (b).

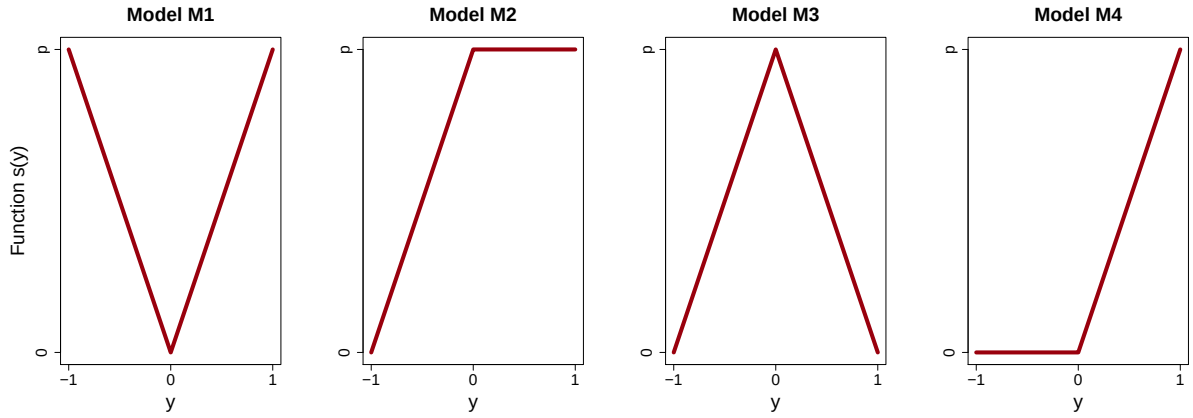


Figure 2: Function $s(y)$ corresponding to simulation models M1-M4.

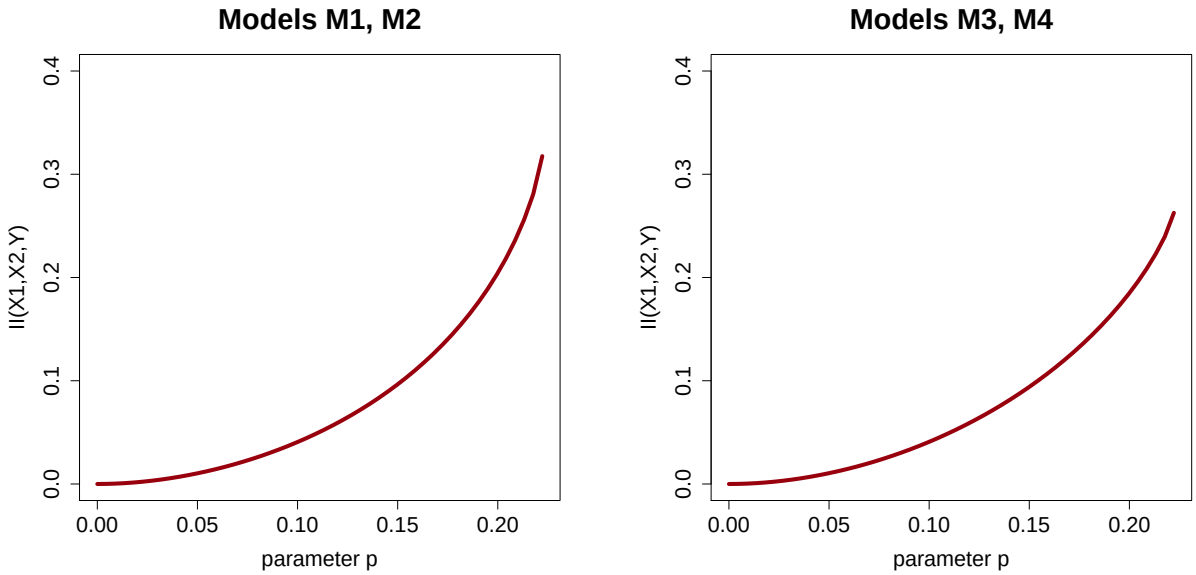


Figure 3: Interaction information with respect to the value of parameter p for simulation models M1-M4

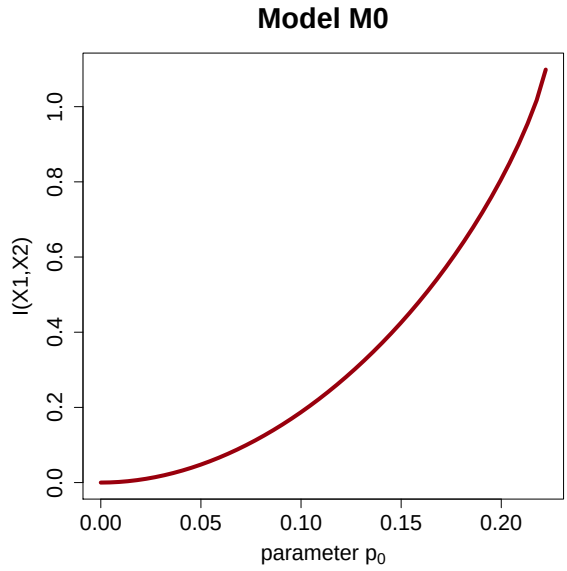


Figure 4: Mutual information with respect to the value of parameter p_0 for the simulation models M0.

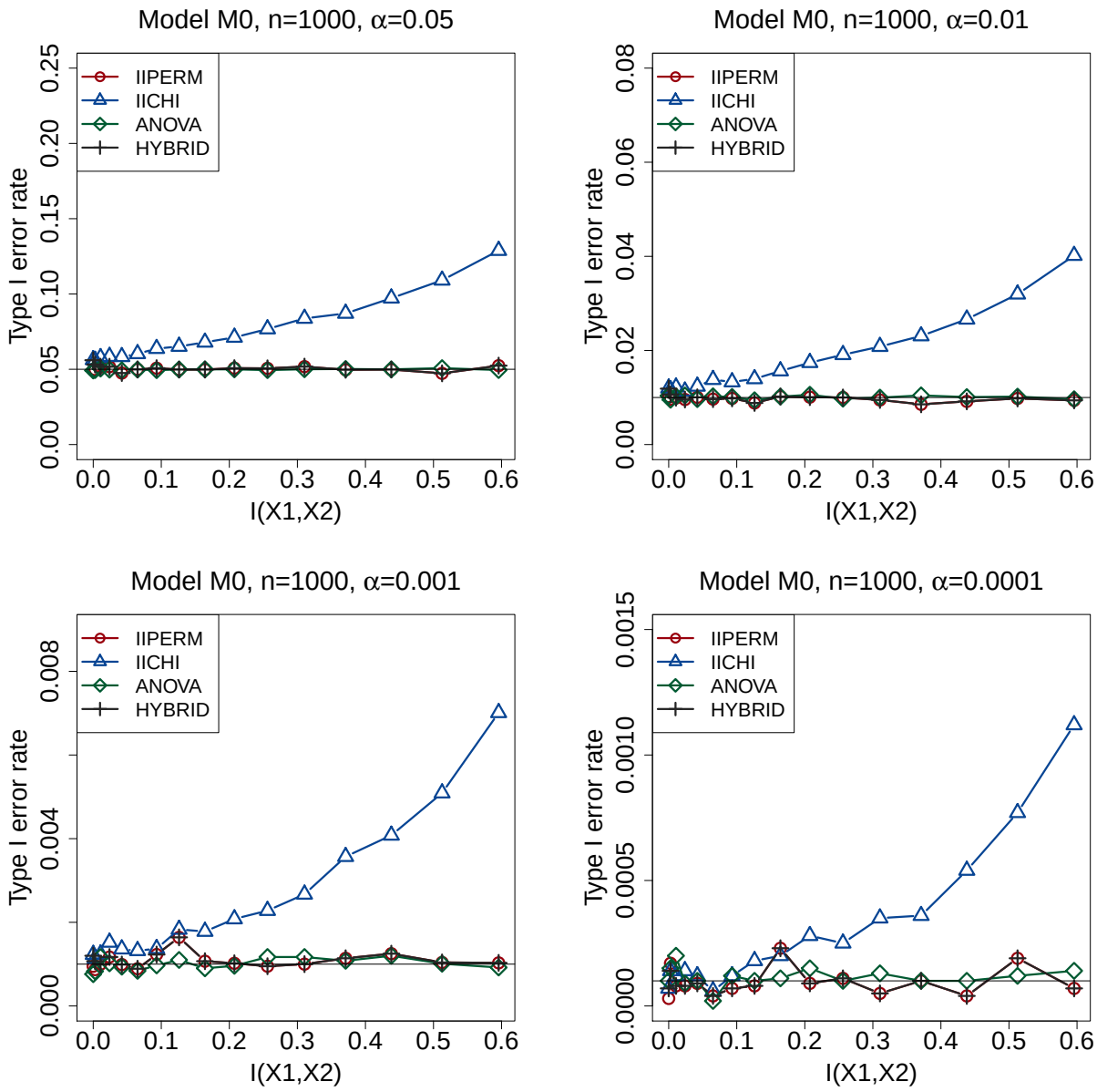


Figure 5: Type I error rate with respect to the mutual information for the simulation model M0, for $\alpha = 0.05, 0.01, 0.001, 0.0001$ and $n = 1000$.

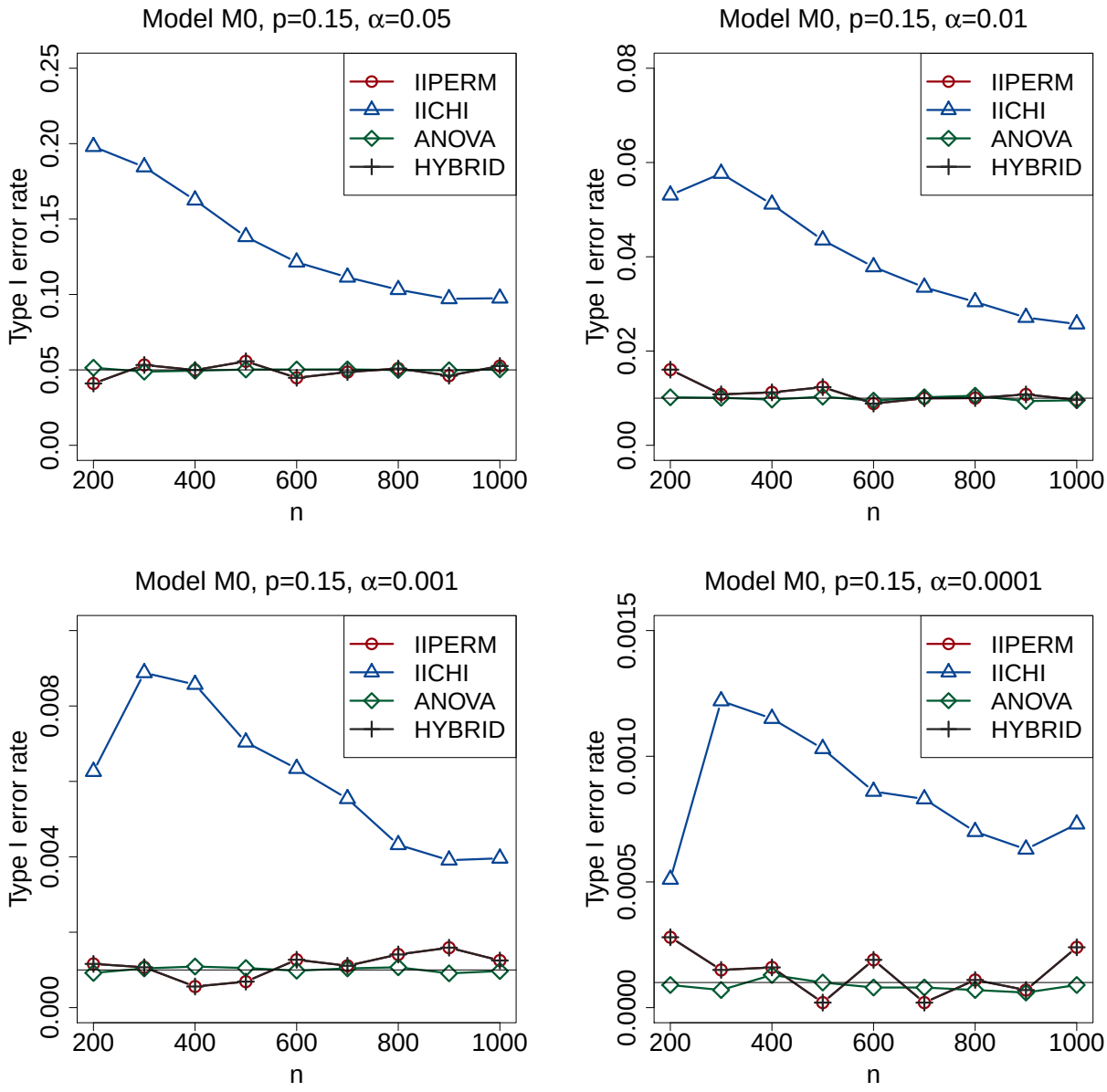


Figure 6: Type I error rate with respect to sample size n for simulation model M0, for $\alpha = 0.05, 0.01, 0.001, 0.0001$ and $p = 0.15$.

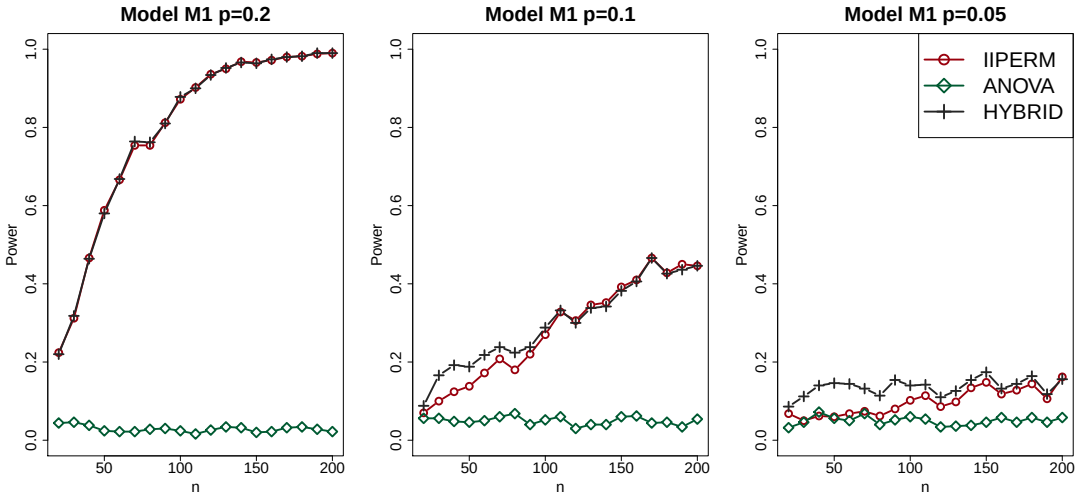


Figure 7: Power with respect to the sample size n for a simulation model M1. $II(X_1, X_2, Y) = 0.1891, 0.0368, 0.0085$, for $p = 0.2, 0.1, 0.05$, respectively.

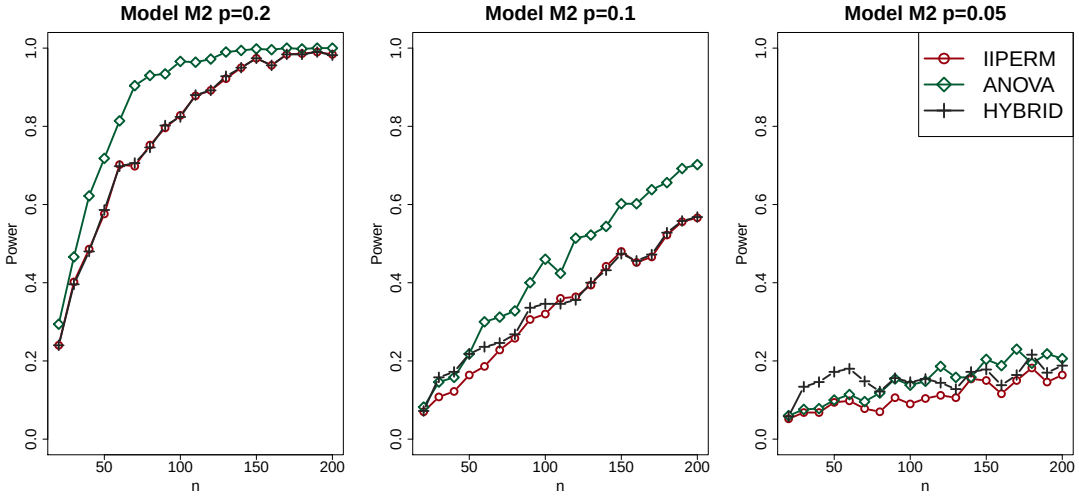


Figure 8: Power with respect to the sample size n for a simulation model M2. $II(X_1, X_2, Y) = 0.1891, 0.0368, 0.0085$, for $p = 0.2, 0.1, 0.05$, respectively.

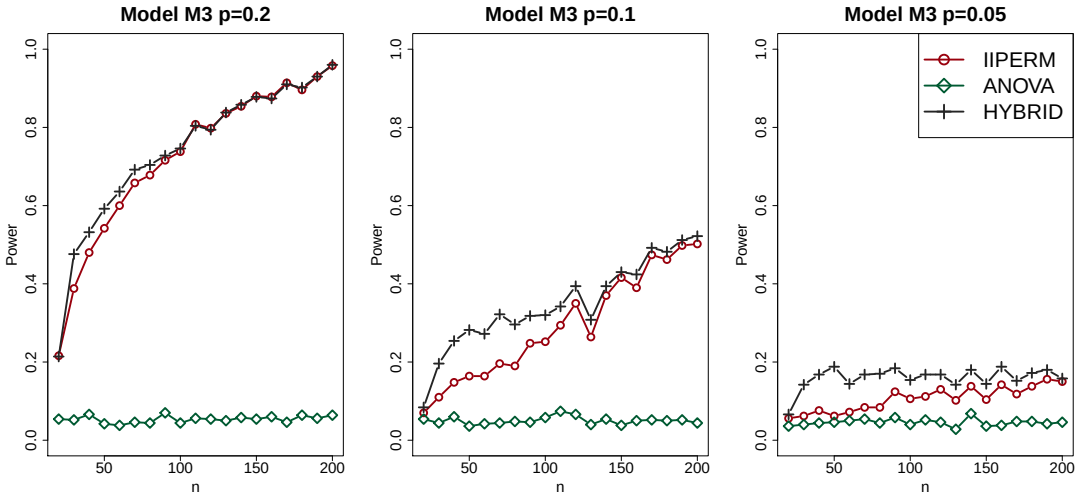


Figure 9: Power with respect to the sample size n for a simulation model M3. $II(X_1, X_2, Y) = 0.173, 0.0371, 0.0086$, for $p = 0.2, 0.1, 0.05$, respectively.

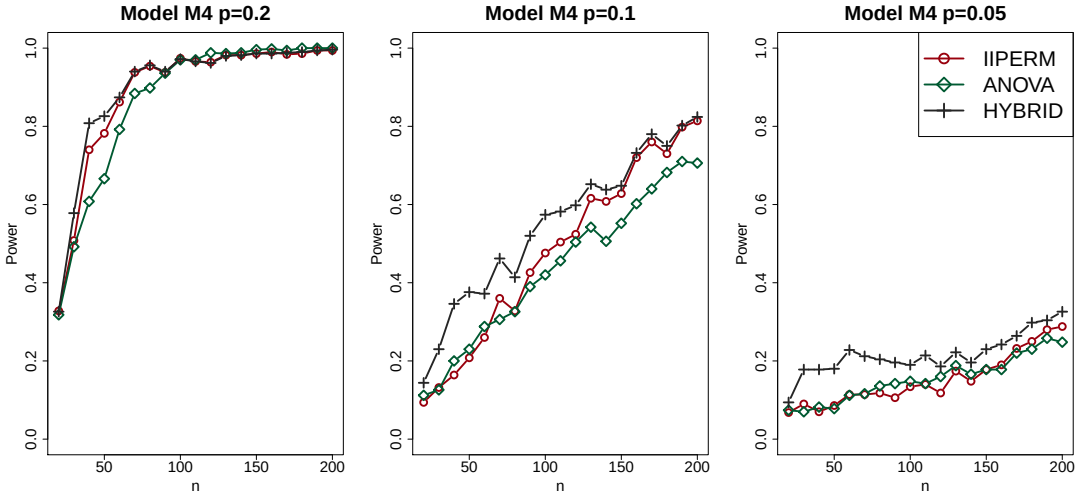
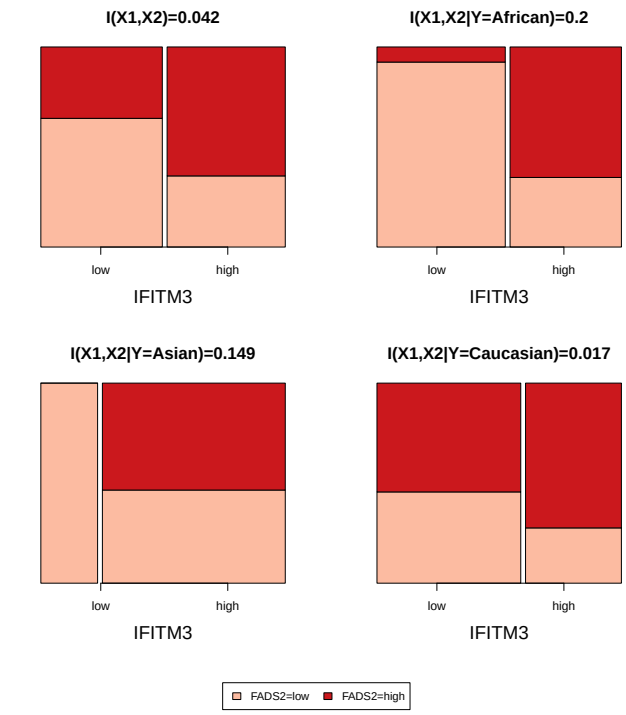
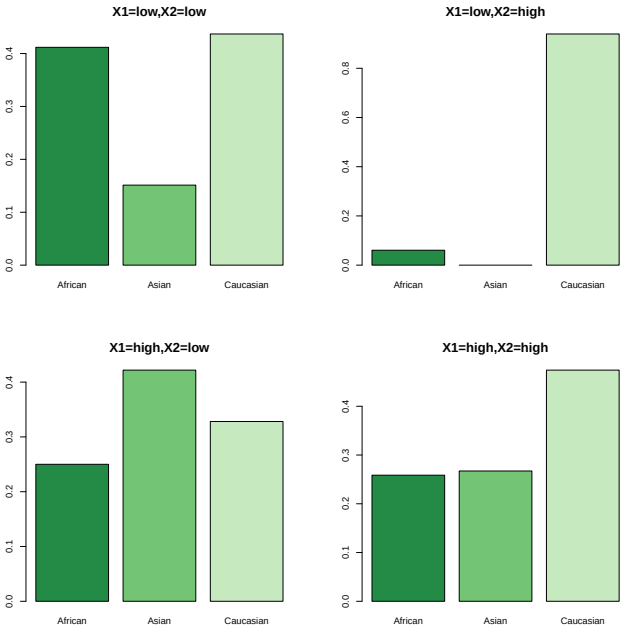


Figure 10: Power with respect to the sample size n for a simulation model M4. $II(X_1, X_2, Y) = 0.173, 0.0371, 0.0086$, for $p = 0.2, 0.1, 0.05$, respectively.

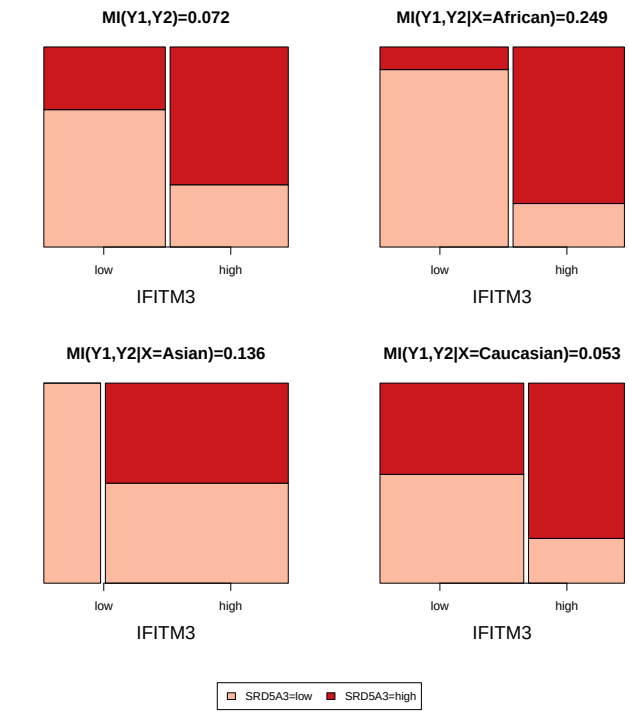


(a)

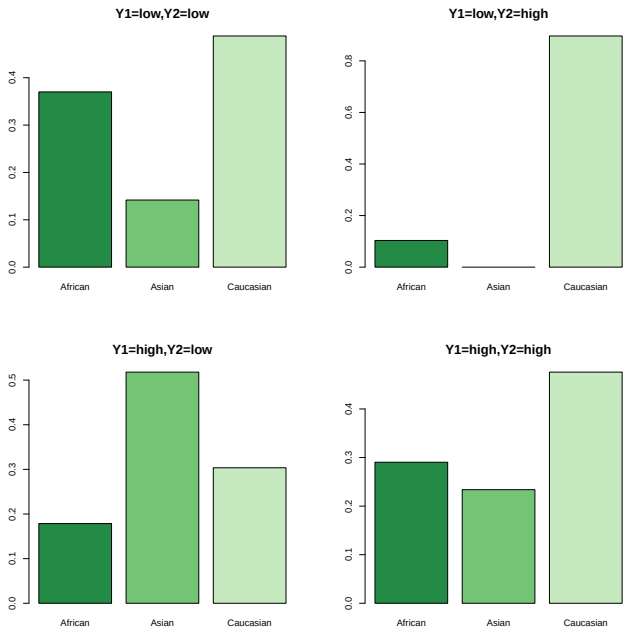


(b)

Figure 11: CD4+ data set. (a) Joint and conditional probabilities for the pair (X_1, X_2)=(*IFITM3*,*FADS2*). (b) Distribution of ethnicity for four combinations of expression levels of genes (X_1, X_2)=(*IFITM3*,*FADS2*).

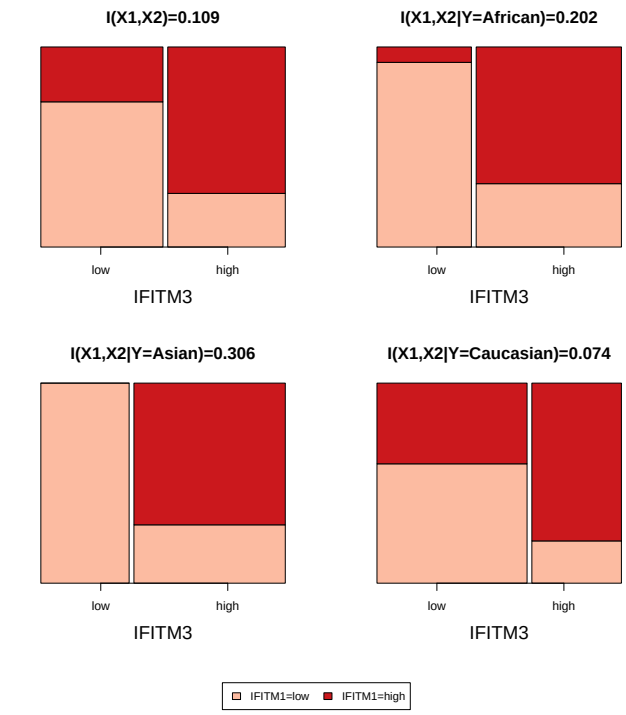


(a)

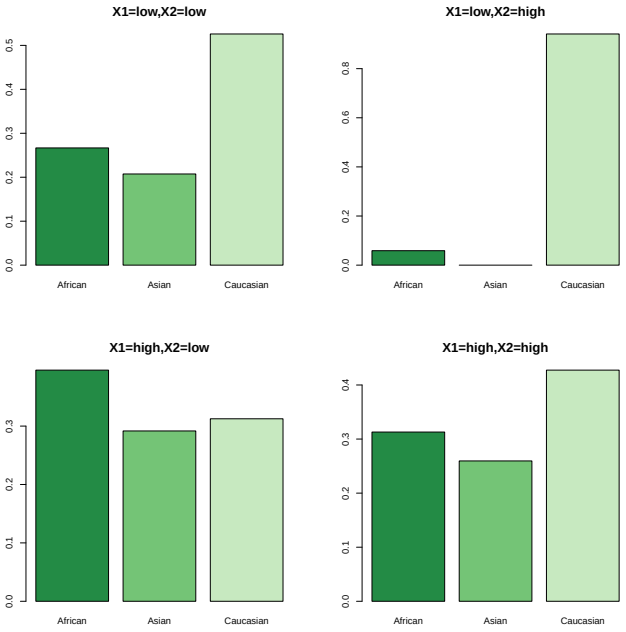


(b)

Figure 12: CD4+ data set. (a) Joint and conditional probabilities for the pair (X_1, X_2)=($IFITM3$, $SRD5A3$). (b) Distribution of ethnicity for four combinations of expression levels of genes (X_1, X_2)=($IFITM3$, $SRD5A3$).



(a)



(b)

Figure 13: CD4+ data set. (a) Joint and conditional probabilities for the pair (X_1, X_2)=(*IFITM3*, *IFITM1*). (b) Distribution of ethnicity for four combinations of expression levels of genes (X_1, X_2)=(*IFITM3*, *IFITM1*).