

Model regresji liniowej dla $X_{1000 \times 950}$ z elementami z rozkładu $N(0, 0.1)$ oraz β z 30 pierwszymi wyrazami równymi 10 i resztą równą 0 oraz σ z niezależnymi elementami z rozkładu wykładniczego z parametrem $\lambda = 1$ oraz z rozkładu Cauchyego.

Użyte kryteria to standardowe mBIC i mBIC2 oraz zmodyfikowane rBIC i rBIC2. W podsumowaniu 100 eksperymentów, zawartych w tabelkach poniżej, widać, że dla modelu gdzie szum ma rozkład wykładniczy nie ma to wpływu, natomiast dla drugiego modelu ogromny, ponieważ klasyczne kryteria prawie nie znajdowały odkryć, mimo dużych wartości parametrów, a zmodyfikowane prawie wszystkie. Widać też, że mBIC i rBIC raczej nie przyjmują fałszywych odkryć, gdy mBIC2 i rBIC2 zdarza się to nieznacznie częściej, lecz przy lepszych wynikach dla prawdziwych odkryć. FDR zazwyczaj wychodzi bardzo mały, a w przypadku nieznaidowania odkryć w drugim modelu wykazuje wartości NaN, przez brak odkryć, zatem FDR=0. Moc w niektórych przypadkach jest bardzo bliska 1, zatem kryteria bardzo dobrze działają.

```
## [1] "exponential"

##           mBIC  mBIC2    rBIC   rBIC2
## prawdziwe 30.0000 30.0000 30.0000 30.0000
## fałszywe  0.0250  1.1750  0.0500  1.5500
## FDR        0.0008  0.0364  0.0016  0.0482
## Moc         1.0000  1.0000  1.0000  1.0000

## [1] "cauchy"

##           mBIC  mBIC2    rBIC   rBIC2
## prawdziwe 0.0750 0.2750 26.5750 29.2250
## fałszywe  0.0500 0.1000  0.0250  1.0500
## FDR        NaN    NaN    0.0009  0.0336
## Moc         0.0025 0.0092  0.8858  0.9742
```

Estymację odporną stosuje się w przypadku, gdy dane pochodzą z rozkładów o ciężkich ogonach oraz gdy mamy dużo obserwacji odstających. Używa się np: funkcji Hubera lub Bisquare, w przypadku takiej estymacji. Po użyciu ich dla rBIC2 MSE zmalało.

Używając tej estymacji w modelach z

Używając zmiennych istotnych wybranych przez rBIC2, wyestymuję parametry metodą najmniejszych kwadratów oraz używając funkcji ...