

Metody oparte na logicznej regresji w zastosowaniu do wykrywania interakcji SNPów

Magdalena Malina

Instytut Matematyczny,
Uniwersytet Wrocławski
pl. Grunwaldzki 2/4, 50-384 Wrocław
e-mail: malina@math.uni.wroc.pl

Małgorzata Bogdan

Instytut Matematyki i Informatyki,
Politechnika Wrocławska
ul. Janiszewskiego 14, Wrocław
e-mail: Malgorzata.Bogdan@pwr.wroc.pl

Rozważamy biologiczny problem detekcji genów odpowiedzialnych za cechy ilościowe. Dane które rozważamy to SNP-y - polimorfizmy pojedynczego nukleotydu, t.j różnice w sekwencji DNA, polegające na zmianie pojedynczego nukleotydu (A, T, C lub G) pomiędzy osobnikami danego gatunku lub drugim, odpowiadającym chromosomem danego osobnika. W tym kontekście niezwykle ważne jest badanie interakcji wielu SNPów, które mogą mieć wpływ na badaną cechę (na przykład status choroby).

Logiczna regresja jest metodą regresji w której predyktorami są logiczne kombinacje zmiennych binarnych. Istnieje kilka wersji logicznej regresji: klasyczna, z algorytmem *simulated annealing* jako metodą przeszukiwania, (Schwender (2007)), logiczna regresja z algorytmem Monte Carlo, zaproponowana przez Kooperberg, Ruczinski czy w pełni bayesowska wersja logicznej regresji (Fritsch, Ickstadt (2007)). Każda z tych metod znajduje kombinacje predyktorów mających wpływ na zmienną objaśnianą i może być zastosowana do danych genetycznych. Porównamy te metody z klasyczną metodą regresji liniowej w zastosowaniu do detekcji interakcji SNPów. Sformułujemy również twierdzenia dotyczące oszacowań błędu pierwszego rodzaju dla tych metod.

Literatura

- [1] Schwender H., *Statistical analysis of genotype and gene expression data*, PhD thesis, URL: hdl.handle.net/2003/23306
- [2] Ruczinski I., Kooperberg C., LeBlanc M., *Logic regression*, J. Comput. Graphical Statist. 12 (3),(2003),474-511
- [3] Kooperberg C., Ruczinski I., *Identifying Interacting SNPs Using Monte Carlo Logic Regression*, Genetic Epidemiology 28, 157-170 (2005)
- [4] Fritsch A., Ickstadt K., *Comparing Logic Regression Based Methods for Identifying SNP Interactions*, Springer Berlin / Heidelberg, Lecture Notes in Computer Science, Volume 4414/2007, pp 90-103