



Mathematics > Probability

Statistical methods of SNP data analysis with applications

Alexander Bulinski (LPMA), Oleg Butkovsky, Alexey Shashkin,
Pavel Yaskov (MIRAS)

(Submitted on 24 Jun 2011)

Various statistical methods important for genetic analysis are considered and developed. Namely, we concentrate on the multifactor dimensionality reduction, logic regression, random forests and stochastic gradient boosting. These methods and their new modifications, e.g., the MDR method with "independent rule", are used to study the risk of complex diseases such as cardiovascular ones. The roles of certain combinations of single nucleotide polymorphisms and external risk factors are examined. To perform the data analysis concerning the ischemic heart disease and myocardial infarction the supercomputer SKIF "Chebyshev" of the Lomonosov Moscow State University was employed.

Subjects: **Probability (math.PR)**; Applications (stat.AP)

Cite as: [arXiv:1106.4989](#) [math.PR]

(or [arXiv:1106.4989v1](#) [math.PR] for this version)

Submission history

From: Alexander Bulinski [[view email](#)]

[v1] Fri, 24 Jun 2011 15:05:57 GMT (247kb)

[Which authors of this paper are endorsers?](#)

Link back to: [arXiv](#), [form interface](#), [contact](#).

Download:

- [PDF](#)
- [PostScript](#)
- [Other formats](#)

Current browse context:

math.PR

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1106](#)

Change to browse by:

[math](#)

[stat](#)

[stat.AP](#)

References & Citations

- [NASA ADS](#)

Bookmark([what is this?](#))

