E. Kochnova[1], I. Shevtsova[2], T. Sobolevsky[1], G. Rodchenkov[1]

Software for analysis of urinary and blood parameters using Bayesian approach

[1] - Antidoping Centre, Moscow, Russia

[2] - Faculty of Computational Mathematics and Cybernetics, Moscow State University, Russia

Introduction

Research on biological markers is important and progressing field of the biomedical science. From the view point of antidoping laboratories the development of doping markers like testosterone to epitestosterone ratio and hemoglobin (Hb) is now based on population data or longitudinal studies of the same individual. Sottas [1,2] introduced Bayesian analysis to improve the reliability of doping tests. The proposed algorithm allows switching from the population-based intervals to individual ones even when the personal information counts only 1-5 measurements. As personal information accumulates, the population-based intervals turn into the personal ones which may substantially differ for various individuals.

Model

The Model used for software was analogous to that described by Sottas in [2].

Modules

All modules used in this program that were developed by us are presented in the Table 1.

Table 1. The list of developed modules.

Name of module	Type	Purpose	Platform	Development environment
BayesianBiomarkersGui	Executable file	User interface	Microsoft .NET Framework 3.5 SP1	Microsoft Visual C# 2008 SP1
BayesianBiomarkersLib	Dynamic Linked Library	Mathematics software model (calculation of	Windows	Microsoft Visual C++ 2008 SP1

		probabilities and confidence intervals)		
EmfToEps	Dynamic Linked Library	Conversion from Windows Enhanced Metafile to Adobe Encapsulated PostScript	Windows	Microsoft Visual C++ 2008 SP1

All third party modules used in this program are presented in the Table 2.

Table 2. Parameters of third party design modules.

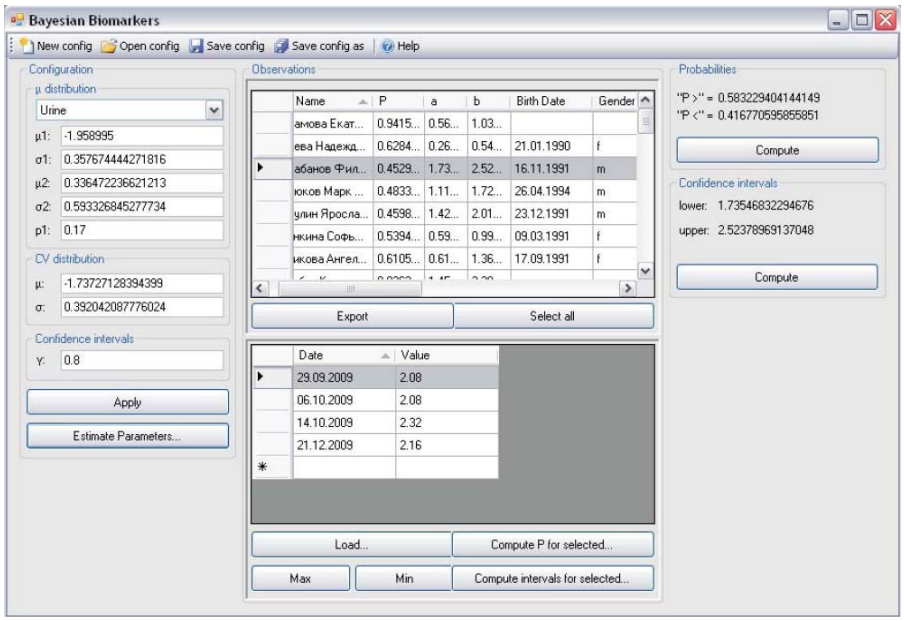
Name of module	Type	Purpose	Version	Platform
GNU Scientific Library	Statistically Linked Library	Numerical integration, searching of equation roots	1.8	Windows
Intel OpenCV	Dynamic Linked Library	Determination of parameters of two normal distributions	2.0	Windows
Microsoft Chart Controls	Assembly	Plotting	1.0	Microsoft .NET Framework 3.5 SP1
LyX Metafile to EPS Converter	Printers driver	Virtual Adobe Enhanced PostScript printer	1.0	Windows

Our goal was to develop a software which allows:

1) to predict the distribution of the forthcoming measurement X_{n+1} given the precedent X_1, …, X_n measurements of urine/blood of an individual are available;
2) to estimate the prior distribution, which is defined from the population, and use the obtained estimate for further prediction. This feature makes our software very flexible, since the user can define the population himself and also select a model for the prior distribution: usually, this is a two-components mixture of log-normal distributions for the parameter μ of urine and a usual (one-component) log-normal distribution for blood. However, one can also use more complicated two-component model for blood, making it possible to mix both male and female data.

For the splitting of Gauss distributions we used a so called expectation-maximization (EM) algorithm which is very effective for the solution of such problems.

Software description



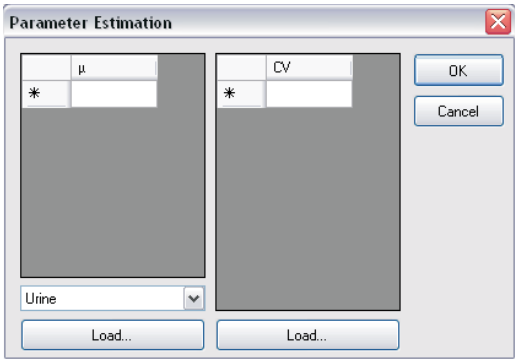
Picture 1. The software interface.

Configuration module

This module allows choosing parameters of distribution from the file or estimate new parameters from distribution files. Files with parameters of distribution could be saved with appropriate names and used later.

Confidence intervals

Presented as level of confidence for the calculations of intervals.



Picture 2. Parameter estimation module.

Parameter estimation

This module allows loading or entering data of population. Program automatically calculates parameters of a new distribution (see pic.2).

Observation module

This module shows full information about athlete, dates and results of the analysis. For each athlete conditional probability (P) and confidence interval (lower - a and upper - b) could be calculated. Observations could be loaded from file or entered manually. Program allows calculating parameters for a number of athletes and export the results into excel file.

Results of estimation of prior probability distributions with an application to T/E ratio

To estimate the prior distribution of Russian athletes 1764 samples (1083 males and 681 – females) were analyzed during 2008-2009 years. T/E data for men's and women's populations and also for combined population were obtained. All three distributions are presented in Fig.1.

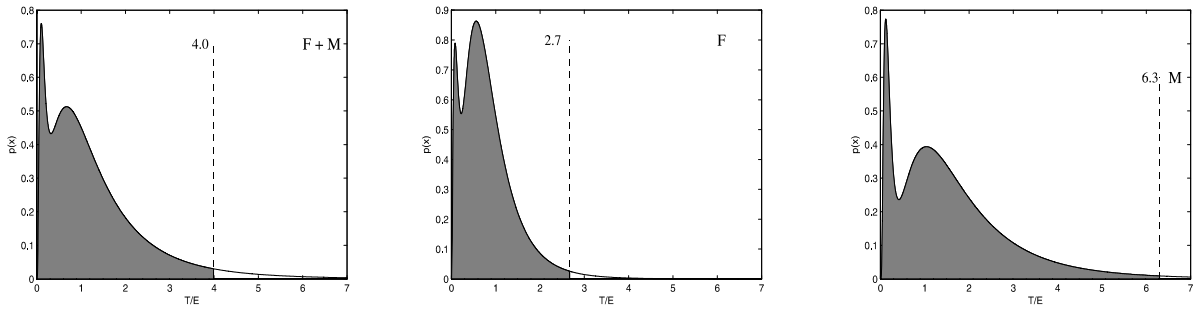


Figure 1. The prior distribution of μ is taken as the sum of two log-normal distributions for Male+Female (M+F), Female (F) and Male (M) populations. Dashed lines are the upper limits (p=0.95).

Conclusion

Developed software allows estimating different prior distributions depending on gender, age, race, sport etc. Our data demonstrated that gender-dependent limits should be used to achieve a more reliable analysis of endogenous steroids.

References

1. Sottas PE, Baume N, Saudan C, Schweizer C, Kamber M, Saugy M. Bayesian detection of abnormal values in longitudinal biomarkers with an application to T/E ratio. Biostatistics. 2007 (8) p. 285-96.
2. Sottas PE, Saudan C, Schweizer C, Baume N, Mangin P, Saugy M. From population- to subject-based limits of T/E ratio to detect testosterone abuse in elite sports. Forensic Sci Int. 2008 174 (2-3), p. 166-172