

Application of Bayesian statistics in social sciences

Annotation

There are two paradigms in contemporary statistics. The first one, which is the best known and widely used, is called “frequentist”. The second one is practically unknown (especially in Bulgaria) and it is used by limited group of scientists (mainly in the world). It is called Bayesian.

The goal of the course is to present theoretical introduction as well as demonstration of the possible applications of Bayesian statistics in the field of the social sciences.

The principles, instruments and techniques of Bayesian statistics will be presented in the theoretical introduction. This means that there necessary will be a little bit more mathematics in this section. However, it is necessary base when various applications are examined. But only such kind of practical problems, which frequentist statistics can not solve (or at least, can solve with serious difficulties), will be examined in the course. They are:

- Inference based on information from non-representative samples. Lately there are enough examples about non-representative samples – votes on Internet, votes by phone or by SMS concerning various TV shows (Bulgarian Song for Eurovision, Test of the Nation, Big Brother, Star Academy, Musical Ku-ku Academy, Vote of Thrust, Crash etc.).
- Inference based on information from representative samples but concerns to several parameters which are not independent. For example, the problem about estimation of relative frequency of people who choose certain party only among the actual voters.
- Calculation of probability for belonging to certain group. This can be successfully combined with multiple tables. This way easy alternative of methods called either “risk analysis” or “decision making under uncertainty” is received.

Syllabus

1. A little bit history

Probability theory.

Frequentist statistics.

Bayesian statistics.

2. A little bit mathematics

Probability. Conditional probability.

Sum rule and product rule. Marginal probability and Bayes theorem. Prior probability and posterior probability. Method of maximum entropy.

Normalization and marginalization. Elimination of nuisance parameters.

Distribution, probability density and probability distribution function.

3. Statistical inference of relative frequencies based on prior probability

Sample without return, sample with return, large population.

Calculation of probabilities.

Probability distribution, hypotheses testing and confidence intervals.

4. Statistical inference of important parameters of distribution (mean, standard deviation etc.) based on prior probability

Sample without return, sample with return, large population.

Calculation of probabilities.

Probability distribution, hypotheses testing and confidence intervals.

5. Application in the field of social sciences – non-representative samples

Votes on Internet.

Votes on TV – Bulgarian Song for Eurovision, Big Brother, Star Academy, Vote of Thrust, Crash, Test of the Nation, Musical Ku-ku Academy etc. The problem about the determination of the winner.

Students' evaluation of teaching.

6. Statistical inference of relative frequencies based on posterior probability

Sample without return, sample with return, large population.

Calculation of probability.

Probability distribution, hypotheses testing and confidence intervals.

Most probable distribution (most probable relative frequencies) in the population.

7. Statistical inference of important parameters of distribution (mean, standard deviation etc.) based on posterior probability

Sample without return, sample with return, large population.

Calculation of probability.

Probability distribution, hypotheses testing and confidence intervals.

8. Application in the field of social sciences – representative samples

Estimation of two and more relative frequencies which are not independent.

Electoral surveys. The problem about relative frequency among the actual voters.

9. Estimation of probability for belonging to certain group

Calculation of probability for belonging to certain group based on prior and posterior probabilities.

Bank rules for granting a loan. Insurance rule for setting insurance premium.
Possibilities for application in the field of marketing and political surveys.

10. Application in the field of forecasting – confidence interval of extrapolations

Analysis of residuals and calculation of probability for belonging to certain group based on prior and posterior probabilities.

Confidence intervals of extrapolations.

Bibliography:

- Bretthorst, L.** 1990. An Introduction of Parameter Estimation Using Bayesian Probability, in Maximum Entropy and Bayesian Methods, P. Fougere (ed.), Kluwer Academic Publishers, Dordrecht the Netherlands
- Dose, V.** 2002. Bayes in Five Days, Lecture notes from a ten hour tutorial on Bayesian analysis given at the International Max-Planck Research School on bounded plasma
- Haralampiev, K.** 2006. Bayesian Inference of Relative Frequency (In the Case of Electoral surveys), Annuaire de l'universite de Sofia "St. Kliment Ohridski", Fakulte de philosophie, Livre – Sociologie, Tome 99
- Jaynes, E.** 1986. Bayesian Methods: General Background, in Maximum-Entropy and Bayesian Methods in Applied Statistics, J. H. Justice (ed.), Cambridge University Press, Cambridge
- Jaynes, E.** 1988. The Relation of Bayesian and Maximum Entropy Methods, in Maximum-Entropy and Bayesian Methods in Science and Engineering, Vol. 1, G. J. Erickson and C. R. Smith (eds.), Kluwer, Dordrecht
- Jaynes, E.** 2003. Probability Theory: the Logic of Science, Cambridge University Press, Cambridge
- Loredo, T.** 1990. From Laplace To SN 1987A: Bayesian Inference In Astrophysics, in Maximum Entropy and Bayesian Methods, P. F. Fougere (ed), Kluwer Academic Publishers, Dordrecht
- Харалампиев, К.** 2004а. Анкетите в интернет: възможност за статистически изводи и интерпретиране на резултатите, Социологически проблеми, брой 3-4
- Харалампиев, К.** 2004б. Нетрадиционен поглед върху традиционни статистически проблеми, Балкани, София
- Харалампиев, К.** 2005. Телевизионните гласувания по телефона – проблемът за „победителя”, Социологически проблеми, брой 3-4
- Харалампиев, К.** 2006. Студентската оценка на преподаването – проблемът за точността на изводите, Юбилеен сборник „30 години катедра „Социология””