A Parallel Discussion of Classical and Bayesian Ways as Introduction to Statistical Inference

- Teacher training in Hungary

Ödön Vancsó
Mathematics Education Center of Eötvös Lóránd University Budapest (Hungary)
12.07.2008 Monterrey ICME-11 TSG-13

The program established in 2003

- I have offered a two-semester seminar for pre-service teacher students since 2003 (lower & higher secondary level teachers; 2 hrs per week).
- Yearly about 10 students have taken part in this seminar (numbers ranging from 4 to 19).
- The program has changed several times in details but not conceptually.
- The course is optional for students.
- You can read more about the content in my paper.

First semester

- Conditional probability and probability are discussed in connection to real problems.
- We clearly differentiate between the so-called “objective” probability notion and the subjective or subjectivist view on probability.
- The „favourable relation” is discussed by the students (it was introduced by Chung 1942). This relation is a weak form of implication.

Second semester

- Introduction to statistics is done parallel in the two theories using such examples, which are suitable for both methods.
- History of inferential statistical is presented as an example how two different aspects of problem solving can be developed.
- Interpretation of the results from classical and Bayesian analysis.

Brief survey of the topic

- Inferential statistics started in the 1920's & 30's

  The main persons are on the classical side:
  R.A. Fisher, J. Neyman, K. Pearson

  and the Bayesian side:
  B. de Finetti, D. Lindley and J. Savage

The central classical notions and ideas

- Hypothesis test (Fisher)
- Confidence interval (Neyman)

Based on a frequentist notion of probability bound to a „true” probability situation, i.e. a situation repeatable under the same conditions

Interpretation

- What does the confidence interval really mean?

I collected data
about it in 2004 via the Internet.
I asked only people who had studied mathematics
at least 3 years at a university before.

The result is surprising:
only 2% of the sample could interpret correctly
what a confidence interval means.

The Bayesian way

- The Region of Highest Density (RHD)

  We use an other (wider) probability notion and our status of knowledge will be expressed by a so called prior distribution.

  After collecting data we revise our distribution calculating a new posterior distribution using the Bayes theorem.

An example

Consider the following situation:

Somebody, who does not know how many balls there are in the oldest lottery, arrives in Hungary.

His question is the following:
How many balls are in the box from which the numbered balls are taken?

He knows the result of one lottery (the selected numbers).

Classical idea

- Confidence interval for the maximum number of balls – e. g. for 95% level.

  It has only a „probabilistic meaning” if we apply it many times and then in 95% of all cases the calculated interval contains the maximum number.

  In one single case we *can say nothing* about how probable it is that this interval contains this parameter.

The Bayesian idea

- It gives a direct answer to our question:
  the 95% RHD is an interval which contains the maximum number of balls with probability 0,95.
- But we have „to pay” for this answer
  as we have to use “our” prior distribution,
  which is somehow subjective.

My assumptions

If somebody learns both theories of inference statistics s/he can

- understand better the classical notions at least;
- interpret correctly the result depending on which theory is used
- understand deeper the role of a model
- get a clearer insight into modern maths

Students' comment in interviews

"I understood the confidence interval first after I had become more familiar with the Bayesian region of highest density (RHD)." (student in 2004)

"I really like the Bayesian method because I saw for the first time why people have different opinions in many cases. It is because they have different prior distributions." (student in 2007)

Plans for the future

- Working out some tests and questionnaires, which are suitable to measure the „probability” of my assumptions.
- Gathering more problems and exercises, which can be handled by both methods.
- Working on an implementation of these ideas into secondary school mathematics.

Correction

In the introduction to my article,
D. Wickmann seems to be classified as a Bayesian.

This is not appropriate as he states in a letter to me. He askes me to mention his newest publication which supports his statement.

D. Wickmann:

Stochastik in der Schule 2007

My working place

Eötvös Loránd Tudományegyetem

Természettudományi Kar

Thank you for your attention!

- Contact:

Ödön Vancsó
vancso@ludens.elte.hu