Teaching Statistics in an Activity Encouraging Format

Sytse Knypstra
University of Groningen, The Netherlands

Journal of Statistics Education Volume 17, Number 2 (2009), www.amstat.org/publications/jse/v17n2/knypstra.html

---

Key Words: Cooperative Learning; Peer Teaching; Higher Education; Bachelor Study; Econometrics; Small Groups.

Abstract

In a statistics course for bachelor students in econometrics a new format was adopted in which students were encouraged to study more actively and in which cooperative learning and peer teaching was implemented. Students had to work in groups of two or three students where each group had to perform certain tasks. One of these tasks was: explaining theory and/or solutions of problems to the other groups. In order to prepare them for this task the groups had separate regular meetings with the teacher. Students report higher involvement and greater satisfaction in this format than in the traditional format. For the teacher the format may be more time consuming, but also more rewarding.

1. Introduction

The theory of statistics is an important subject in the study of econometrics at the University of Groningen. During their first year students study the basic concepts of probability theory and inferential statistics. During the first half semester of the second year a course is given focusing on likelihood and asymptotics. For this course the textbook is Azzalini (1996) (chapters 1-4). In this paper we discuss a new teaching format that was adopted for this course.

Until the year 2002 the course was taught in the format of lecture and one tutorial meeting per week. In the tutorial meetings, the students were doing exercises with the help of the teacher and/or undergraduate teaching assistants. This is more or less the standard format in our faculty.

My usual experience was that students were not very active during the course. Many students postponed studying the material until the time the exam was emerging on the horizon. Then time was lacking to study the material in depth, which resulted in unsatisfactory results: students had learned the material only superficially and grades were low.

For the year 2003/2004 I reconsidered the format in order to motivate the students to work harder *during* the course and also to engage in a deeper approach to learning.

2. Considering a new format

For the construction of a new format the following considerations had to be taken into account.

Prerequisites. It was assumed that the students had mastered the concepts of probability theory and statistics they studied in the previous year.

Goals. During the course, the following topics should be covered: sufficiency, exponential families, maximum likelihood estimation, information, asymptotics, likelihood ratio test, Wald test, score test, confidence intervals. After this course, students should be able to recognise and apply these concepts in specific cases. It would be welcomed if in addition they would be able to reflect on the importance and the interrelationships of these concepts.

Conditions. These included the number of students (approximately 30), the availability of one teacher, scheduled time for both the teacher and the students.

Leading principles. There is agreement in the literature (Biggs 2003) that being active while learning and getting regular feedback stimulates the learning process.

0. Therefore the students should be encouraged to work *actively*. Possible activities include making summaries, concept maps and schemes, discussing the subject with the teacher or with other students, answering questions, formulating questions, doing exercises, searching and comparing relevant literature.
a. Students should regularly get feedback on their work.

   At our faculty it is considered important that students learn certain skills:

b. oral presentation skills,
c. written presentation skills,
d. self management,
e. cooperation/team work.

   And finally, I had my own objectives:

f. encourage a critical attitude in students,
g. use cooperative learning and peer teaching in small groups.

The last point (h) needs some further explanation. It is my experience that teaching usually enhances the teacher's understanding. So if you let students explain parts of the theory or solutions to problems to each other then not only the 'receiving' student, but also, and maybe even stronger, the explaining student will benefit. In the literature evidence can be found that supports this view (Garfield 1993; Mazur 1997; Falchikov 2001; Biggs 2003).

In order to implement peer teaching, the 'teacher'-students should have gained a higher level of knowledge of the subject than the 'student'-students. Therefore they have to be instructed separately. This was realised in organising special meetings of the 'teacher'-students with the teacher.

3. An activity encouraging new format

With the guidelines of the previous section in mind, the following format was designed and implemented for the course.

At the introductory meeting, students were asked to form groups of two or three students. In order to keep the format manageable the number of groups (n) should not exceed 10. Each week the same number (n) of assignments was given. In an assignment a certain section of the book had to be summarised or the theory had to be applied to a certain problem, or a problem in the textbook's problem section had to be solved. An example for a typical week is given in Appendix 1. For each assignment four groups had a special role and each group played a special role at four (out of n) assignments.

These roles are:

W. The group that writes a **W**ritten summary of the theory specified in the assignment or a **W**ritten solution to the problem given in the assignment. This amounts to doing the assignment in a pencil-and-paper style. The solution should be well structured and conform to certain guidelines as laid down in a four pages' manual about writing a report and giving an oral presentation.

O. The group that prepares a 10 minutes' **O**ral summary of the specified theory or a verbal exposition of the solution to the problem given in the assignment, similar to a **W**ritten solution. Again, the presentation should be well structured and conform to certain guidelines as laid down in the manual mentioned under **W**. The students may use the traditional blackboard or (powerpoint) slides.

C. The **C**ontact group; this group helps the remaining groups when they encounter problems with the assignment. Once a week each **C**ontact group (separately) has a 20 minutes' meeting with the teacher in the teacher's office where the group's **C** assignment is discussed and obscure points are clarified. This meeting is held in an informal atmosphere, the students are not graded for their contributions during this meeting. The only condition is that they come well prepared: the material should have been studied to a point where the remaining problems can be made explicit. This meeting should suffice to prepare the students in this group for helping the other groups with the assignment.
One hour a week is scheduled for this peer teaching. We will refer to it as the 'noisy session' because all students are simultaneously involved in asking and giving information to their fellow students. Each group splits itself up into one or two members who help the representatives of other groups with their problems regarding the exercise for which the group as a **C** group is reponsible, while at the same time the remaining member(s) ask(s) for clarification of the group's problems with the **W**, **O** and **Q** exercises. Later the gained information should be conveyed to the other member(s) of the **C** group. This is not organised by the teacher, but is the group's own responsibility.

Q. The group formulates at least three **Q**uestions regarding the assignment and also tries to provide answers. It is made clear that questions which force you to reflect on (the background of) the assignment are strongly preferred over simple quiz questions on which only one answer is correct ('What is the formula for the density function of a Gamma distribution'). In case of need for good questions the students may resort to the following examples: 'What is interesting in this exercise?' or 'How does the result change if you change the first assumption into...?' or 'What is the relationship of this problem to surrounding assignments?' or 'Can you find generalisations or interesting special cases?'.
An example of questions and answers given by a certain group in the course in 2007 is given in Appendix 2.

Three plenary sessions are scheduled each week: two sessions for the oral presentations (10 minutes for each group, followed by a discussion of the questions posed by the **Q**-group and possibly questions posed by other students or the teacher). There is a one hour noisy session in which all groups explain problems regarding the assignments to (representatives of) other groups. And once a week the teacher receives each group in his office during 20 minutes' meetings.

All students should be present at each plenary session and at each meeting with the teacher. Absence is allowed at most three times. Although this is seldom necessary, the teacher may mark an insufficiently prepared oral or written contribution or an insufficiently prepared meeting with the teacher as absence. The group's **W**ritten report is graded. In the course a final exam is held; the group grades contribute to the final grade. In this way all leading principles, except for **e** (self management), are implemented:

a. Students are forced to work during the course by doing exercises, making summaries, formulating and answering questions, discussing problems with their peers and studying theory in order to prepare for the meeting with the teacher.

b. Feedback is given by the teacher through correcting the written reports, discussing the questions and answers in class and during the meetings in the teacher's office.

c. and **d.** Oral and written presentation skills are trained when the students have to present their solutions or summaries.

f. Cooperation between members of the groups is necessary in preparing oral and written presentations, questions and answers, and for the meeting with the teacher.

g. A critical attitude and reflection on the relationships of the concepts is encouraged during the meetings with the teacher, by letting the students formulate questions and answers and discussing these in class.

h. Cooperative learning and peer teaching is implemented by letting the students work in groups and letting the contact groups explain solutions to the other groups.

4. Evaluation and discussion

After each course an evaluation is held. Especially interesting in the year 2003/2004 were the answers to the question: *'How do you evaluate this year's format compared to last year's?'* of the 9 students who took the course also in the previous year, in the traditional format. One of these students was rather critical about the new format, in particular about the quality of the oral presentations. The other responses that were obtained were:

- Deeper understanding.
- First I was sceptical about the new method, but it really worked. You have to actively study the material.
- I became more involved. It's a lot more interesting to think for yourself instead of only listening.This year was better. I was forced to really study the material and you don't get everything spoonfed.
- The obligation to do all exercises appeared to be helpful (eventually).
- Good: I was more actively studying; worse: the presentation of the theory.
- It was good to be continuously working on the material, but in this way all meetings were in fact compulsory which is annoying if you have to do other things.
- What I liked better was working in groups.

The answers on the evaluation questionnaire are consistently positive over the years. Generally students report to be more active than in traditional courses. Recurrent critical remarks concern the quality of the presentations by fellow-students and problems in formulating good questions. In order to address these problems feedback is now given on the oral presentations and a number of example questions is supplied (see under **Q**). Still it cannot be expected that the students' presentation skills will match those of a professional and experienced teacher. However, one has to bear in mind that the primary objective of the oral presentations is to provide the presenters with a learning experience.

The new format is suitable not only for statistics courses but for any course where content of a certain theoretical level has to be mastered and where the number of students is at most 30. We recommend that the number of groups should not exceed 10 in order to keep the teacher's time expense limited and group size should not exceed 3 in order to minimize the risk of 'free riding'.

Another necessary condition is that the accompanying text(book) can be studied independently and without too much difficulty. In that case the 20 minutes' meetings with the C-groups will suffice to clarify the obscure points. Ideally all students should study all n exercises, but because no output on the remaining exercises is required, they tend to focus on the four exercises for which their group has a special (**W**, **O**, **C** or **Q**) task. Therefore multiple exercises should cover the most important topics and the teacher should assign exercises to groups in such a way that each group has to deal with these topics in one way or another.

It may seem that the new format is more time consuming for the teacher than the traditional one because the teacher spends n x 20 minutes a week meeting the student groups separately, which he/she doesn't in a traditional setting. However an exact comparison is hard to make because much depends on how the traditional lectures and tutorials are organised and prepared. The teacher will generally need more time to prepare (slides and the organisation of interaction) for a traditional lecture than a plenary session in the new format. Also the correction of individual work, if any, may possibly take more time than correcting the n groups' written work.

Definite advantages of the new format are a higher involvement and more satisfaction of the students, and, as perceived by the teacher, a higher level of understanding, at least in some students. For the teacher it is most rewarding to see students really work, to see that they get interested and involved. And it is a pleasure to talk with them each week about the interesting subject of theoretical statistics in an informal, interactive, small scale setting.

Appendix 1

Assignments for week 1

Background theory: Azzalini sections 2.1 and 2.2.1. Topics: statistical models, parametric models, parametrizations, the likelihood function, the likelihood principle.

The codes behind the problem number indicate the groups which have special tasks regarding the problem. For example for problem 1, group 6 is the contact group, group 2 writes a written solution, group 5 prepares an oral presentation and group 3 formulates (and tries to answer) at least three questions about problem 1.

1. (**C6**, **W2**, **O5**, **Q3**) Summarise section 2.1 (Azzalini).

2. (**C7**, **W8**, **O1**, **Q4**) Summarise section 2.2 (Azzalini).

3. (**C8**, **W3**, **O4**, **Q2**) (Exercise 2.3, Azzalini). Consider an s.r.s. of size *n* from a random variable with density function at *t* given by

$$g(t;\theta) = \begin{cases} e^{-(t-\theta)} & \text{if } t > \theta, \\ 0 & \text{otherwise,} \end{cases}$$

   where *?* ? R.
   (a) What is the sample space?
   (b) Write down the corresponding likelihood function

4. **C2**, **W7**, **O6**, **Q1**) In an urn we have 100 (red and white) balls. Let *?* be the proportion: *number of red balls/total number of balls* in the urn. Balls are drawn from the urn with replacement until a red ball is drawn.
   (a) Find the parameter space in this case.
   (b) What is the probability distribution of the number of draws up to and including the draw where the first red ball appears? Find the expectation of this number of draws.

   This experiment is performed 5 times. Subsequently the numbers of draws are 6, 3, 4, 5 and 7.
   (a) Find the probability of this outcome for the cases *?*= 0.2 and *?* = 0.3.
   Which value of *?* is more likely?
   (b) Find the likelihood function and the maximum likelihood estimator for *?*. Find the maximum likelihood estimate for *?* based on the observations in part **c**.

5. (**C3**, **W6**, **O2**, **Q7**) Azzalini (Example 2.1.3) states:
   A parametric family of distributions is said to form a location and scale family if the density function of a single component can be written as

$$g(t;\theta_1,\theta_2) = \frac{1}{\theta_2} g_0\left(\frac{t-\theta_1}{\theta_2}\right)$$

for a fixed density $g_0(t)$ and two parameters $?_1$, $?_2$, called location and scale parameters, respectively ($?_2 > 0$). The reason for the terminology is that all distributions of the family can be obtained by applying a linear transformation of the form

$$Y = ?_1 + ?_2 Y_0$$

to a variable Y_0 with density $g_0(\cdot)$.
Prove this statement and find examples of location and scale families of distributions.

6. (**C1**, **W5**, **O3**, **Q6**) (Azzalini Example 2.1.4).
What is a Poisson process? What is its link to Poisson- and exponential distributions? Show how the number of events in a certain time interval and the distribution of times between events are related.

7. (**C5**, **W4**, **O7**, **Q8**) (Azzalini Example 2.1.5).
Find an unbiased estimator for s^2 based on an SRS $Y_1, \ldots Y_n$ from a $N(?, s^2)$ distribution. Is the square root of this estimator an unbiased estimator for s? Explain.

8. (**C4**, **W1**, **O8**, **Q5**) Given is the pair of random variables (Y_1, Y_2). Both variables can only take the values 0 and 1. The first variable (Y_1) has a Bernoulli distribution with parameter p and for the second variable (Y_2) the following holds:
$(Y_2|Y_1 = 0) \sim$ Bernoulli(p_0) and $(Y_2|Y_1 = 1) \sim$ Bernoulli(p_1).

(a) Find the sample space.
(b) Find the parameter space.
(c) Show that $\Pr(Y_1 = y_1, Y_2 = y_2) = \pi_{y_1}^{y_2}(1-\pi_{y_1})^{1-y_2}\pi^{y_1}(1-\pi)^{1-y_1}$ for y_1 and $y_2 = 0,1$.
(d) Are Y_1 and Y_2 independent? Are they identically distributed?

Given is a sequence $(Y_1, \ldots Y_n)$ of random variables. Each variable can take only the values 0 and 1. The first variabele (Y_1) has a Bernoulli-distribution with parameter p and for $j = 2, \ldots, n$ the following holds: $(Y_j \mid Y_{j-1} = 0) \sim$ Bernoulli (p_0) and $(Y_j \mid Y_{j-1} = 1) \sim$ Bernoulli (p_1). The Markov-property holds, which means that the distribution of Y_j, conditional on the outcomes of all preceding Y_i's, is identical to the distribution of Y_j, conditional on the outcome of Y_{j-1}.

(a) Find the sample space.
(b) Find the parameter space.
(c) Find the likelihood function.

Appendix 2

Questions and provisional answers to problem 8 in Appendix 1 as given by one group.

Question 1: *What is interesting about this question?*

Answer 1: It is possible to find the likelihood function for random variables that have a conditional distribution.

Question 2: *What are the possible outcomes for the sequence* $y_1, \ldots ,y_n$ *if we know that* $p_0 = 0$ *and* $p_1 = 1$?

Answer 2: There are two possible outcomes depending on the outcome of Y_1. If $Y_1 = 0$, then we will get a sequence of *n* zeros and if $Y_1 = 1$, then we will get a sequence of *n* ones.

Question 3: *If you know the whole sequence can you tell something about the first parameter p?*

Answer 3: Only the first outcome of the sequence tells you something about p since Y_1 is the only random variable in the sequence that is distributed Bernoulli(p). If $y_1 = 0$, we would guess $p < 0.5$ and if $y_1 = 1$, we would guess $p > 0.5$.

The author's comment: Question 1 was given as an example question to the students. The answer, which could have been better phrased, rightly mentions the feature (conditionality) that distinguishes this problem from most other textbook problems. The second question concerns a (trivial) special case. I would have preferred an effort of generalising the problem or looking for related problems or for practical applications. The last question hints at estimating a parameter, but in their answer the group fails to formally derive an estimator (estimation is the subject of chapter 3, to be studied in a later week).

References

Azzalini, A. (1996), *Statistical Inference Based on the Likelihood*. London: Chapman & Hall.

Biggs, J. (2003), *Teaching for Quality Learning at University : What the Student Does* (2nd ed.), Maidenhead: The Society for Research into Higher Education & Open University Press.

Falchikov, N. (2001), *Learning Together: Peer Tutoring in Higher Education*. London: Routledge Falmer.

Garfield, J. (1993), "Teaching Statistics Using Small-Group Cooperative Learning," Journal of Statistics Education, Vol.1, n.1.

Mazur, E. (1997), Peer Instruction: a User's Manual. Upper Saddle River, NJ: Prentice Hall.

Sytse Knypstra
Economics and Management
University of Groningen
The Netherlands
S.Knypstra@rug.nl
