A WORKSHOP IN PAKISTAN ON TEACHING STATISTICAL INFERENCE THROUGH A FREELY AVAILABLE ONLINE SOFTWARE

Saleha Naghmi Habibullah
Kinnaird College For Women, Lahore, Pakistan
salehahabibullah@gmail.com

Statistics-teaching in some of the developing countries is pretty theoretical making it difficult for students to grasp fundamental concepts of statistics. Whereas in the technologically advanced countries there is now great emphasis on computer-based simulations for facilitating understanding of key concepts, numerous academics in the developing world are still not using this approach due to lack of awareness as well as lack of resources. This paper reports on a workshop conducted in Pakistan using a freely available online software in order to introduce faculty-members of various universities and other higher education institutions to the simulation-based method of learning and teaching core concepts of statistical inference. Feedback from participants indicates that this method is not only very enjoyable but also very effective.

INTRODUCTION

No matter how fascinating the discipline of Statistics may be for some, the fact of the matter is that, by and large, students find concepts of statistical inference pretty difficult and daunting. (See, among others, Onwuegbuzie, 1997, Suanpang et al., 2004 and Habibullah, 2014). The problem is *all the more advanced* in many developing countries of the world where the teaching of Statistics is by and large textbook-oriented resulting in a "mechanical style of delivery" on the part of the teachers and rote-learning of formulae, theorems, etc. on the part of the students.

In the technologically advanced countries, during the past few decades, the ability to generate random numbers through computers has led to large-scale simulations that were *unthinkable* in the pre-computer era and which greatly assist in understanding fundamental properties of sampling distributions, confidence intervals and p-values. In the developing world, however, a huge majority of teachers, lecturers and professors is not using computer technology for teaching purposes due to lack of awareness as well as lack of resources. The increasing availability of freely downloadable software on the internet can be regarded as a blessing due to which it has now become possible for academics in the developing world to carry out simulation studies for teaching purposes. The following sections of this paper report on a workshop conducted in Pakistan in order to introduce faculty-members of various universities /higher education institutions to the simulation-based method of teaching fundamental concepts of statistical inference using a freely available online software called StatKey. Feedback obtained from participants is included to enable the reader to assess the usefulness of this exercise.

BACKGROUND

In the summer of 2014, the author attended an Invited Talk in which the speaker demonstrated various aspects of StatKey, a free online technology designed to help introductory students understand and easily implement bootstrap intervals and randomization tests. (See Morgan et al., 2014.) The quality of the talk coupled with the functionality of the software was highly inspiring and motivational, and led the author to decide that she will try to introduce this method of learning and teaching in Pakistan. Due to multifarious teaching duties, research-related responsibilities and administrative initiatives, it was not before the Fall of 2017 when the author was actually able to take the first step in this direction.

PREPARATIONS FOR THE WORKSHOP

In November 2017, the author approached the founder and patron of the Islamic Countries Society of Statistical Sciences (ISOSS) as well as of the Pak Institute of Statistical Training And Research (PISTAR) with the request that she may be allowed to conduct a Workshop on *Teaching Statistical Inference Through a Freely Available Online Software* at ISOSS House, the Society Headquarter, on Sunday, 12th November, 2017. Having obtained the green signal, the author commenced the process of preparation of the workshop flyer which was disseminated among the

ISOSS community by Executive Secretary ISOSS through email, Facebook and WhatsApp. In order to attract *as many* persons to the Workshop as possible, the flyer contained no mention whatsoever of any registration fee for participation in the Workshop. The prospective participants were encouraged to bring with them their internet-enabled laptops so that they would be able to practice the various steps then and there.

Equally importantly, the author prepared a feedback questionnaire consisting of more than a dozen questions. As well, the author prepared slides with the help of the Invited Paper by Morgan et al., 2014 published in the Proceedings of ICOTS-9. The total number of slides turned out to be 166.

CONDUCT OF THE WORKSHOP

The Workshop attracted sixteen persons comprising college/university teachers, researchers and students of MPhil Statistics. About four or five of them were equipped with internet-enabled laptops. At the outset, the resource person (i.e. the author) asked the participants whether they had ever heard about StatKey. Not a single participant indicated that they had any knowledge whatsoever about this software. The resource person then shared with them her experience of the ICOTS9 conference held in Flagstaff, Arizona in July 2014 and, in particular, the Invited Talk by one of the developers of StatKey. The participants were taken to the relevant website and it was pointed out that the software is also available freely as an App which allows it to be used *even without an internet connection.* Electricity shortage and load-shedding being a 'perpetual' problem in developing countries such as Pakistan, this feature of the software was appreciated by the participants.

Throughout the workshop, the resource person alternated between exposition of the slides prepared by her and demonstration of statistical concepts through StatKey. All in all, the participants witnessed/experienced about one and a half hour of productive rendering by the resource person regarding the utilization of the freely available online software for constructing bootstrap confidence intervals, carrying out randomization-based hypothesis-testing, construction of *sampling distributions*, calculation of summary statistics of data-sets as well as graphs of *theoretical distributions* that could be created using the software. In addition to utilization of the data-sets that were embedded in the software, some fictitious data-values were used to demonstrate the software's capability to deal with any data-set that a user might be wanting to work with.

During the workshop, more than once, the resource person emphasized that in Pakistan as well as in other South-Asian countries, although, historically, there has been a great emphasis on mathematical explication of various concepts including those pertaining to statistical inference (and rightly so), for a large majority of the people "out there" who are not statisticians, it is very *difficult* and *tedious* to understand concepts such as sampling distributions, central limit theorem and p-values. She asserted that computer-based simulations such as those possible by this freely available online software are *just the right thing* for making these concepts easier to comprehend and be comfortable with.

USEFULNESS OF THE WORKSHOP

As indicated by the participants during the very first few minutes of the workshop, other than a very few, nobody was familiar with the concepts of bootstrapping, randomization hypothesis-tests or simulation-based procedures in general. Ignorance regarding resampling techniques served as a *plus*! As the resource person went through various sections of the workshop, she witnessed an *ongoing interest* on the part of the participants in the various procedures and concepts that were conveyed (bootstrap intervals followed by randomization tests, descriptive statistics and graphs, data-entry, sampling distributions and theoretical distributions).

As far as the resource person is concerned, it was the first time she was working with this software herself and, as such, she experienced not only development of a better understanding of the concepts of bootstrapping, randomization distributions and randomization tests but also noted that some ideas were generated in her mind regarding "well-understood" concepts in ways that she had never experienced before.

FEEDBACK FROM PARTICIPANTS

Immediately after the culmination of the workshop, the resource person distributed the feedback questionnaire requesting the participants to answer the questions *as honestly as possible*. This section contains a preliminary analysis of the responses received from the participants.

One of the important questions was: 'Which portions of the Workshop were a source of new *information / learning / knowledge* for you?' A compilation of the responses is as follows:

- *The concepts of Simulation and Bootstrapping*: Five out of sixteen i.e. almost one-third of the participants had never heard about simulation and another five had never heard about bootstrapping prior to the workshop;
- *The concept of Randomization Hypothesis Tests*: Eight of the sixteen participants had never heard about randomization tests;
- *The concept of p-value*: Eight participants had *some* idea about p-value but this concept was not very clear to them;
- *The usefulness of transition from simulation to distribution-based methods*: As many as nine i.e. more than one-half of the participants were *not aware* of the usefulness of transition from *simulation* to distribution-based methods prior to the Workshop.

Two of the questions probed the participants' expectations from the workshop *before* they arrived at the venue and *the type of experience the workshop had been* for them. Responses included the following:

- Lecturer and PhD Scholar (Statistics): "I was expecting some new software for statistical computations, with the help of this new software, I can teach Statistics to non-Statistical students like business students"; "it was a great learning experience."
- Medical doctor, researcher, senior citizen: "Scared from new"; "wonderful, exciting."
- MSc Statistics, MBA Marketing, PGD MIS, currently Teaching and Consulting: "To see statistics dynamically working"; "Felt statistics today. Saw statistics in action today."

One of the questions sought the remarks of the participants regarding the *teaching style* of the resource person. Some of the responses were, "Pleasant and easy to understand. Her command on the subject and fluency in the language was impressive", "Awesome and interesting teaching style", "Very nice but slides are little boring."

Toward the end of the questionnaire, suggestions were sought from the participants for improvement of similar Talks/Workshops in the future. A compilation of the suggestions forwarded is as follows:

- The workshop series should be *continued* (i) with various groups of participants including college students, (ii) with the *same* group of participants participating in a number of workshops one after the other.
- *Proper advertisement* and *marketing* is essential in order to attract relevant persons not only from the statistics community but also from other communities such as medical doctors.
- Slides should be *emailed* to registered participants prior to the workshop.
- *Enough time* should be given to such training (more than the amount of time that was allocated for this workshop).
- Venue should have *proper décor* with *comfortable seating* and other relevant arrangements.
- *Two pieces of multimedia* should be available simultaneously, one for the online software and one for the slides.
- The *actual data-set* which users can insert in the software should also be discussed in accordance with the various concepts being taught and learnt.

DISCUSSION AND CONCLUDING REMARKS

In all parts of the world, non-trivial proportions of students find statistical concepts difficult and commit a variety of *errors* which have been the subject of research for the past few

decades. Batanero et al. (1994) present a survey of the reported research about students' *errors*, *difficulties* and *misconceptions* concerning elementary statistical concepts including concepts of statistical inference such as tests of hypotheses. Garfield and Ben-Zvi (2007) review current research on teaching and learning statistics and focus on questions such as 'What are some of the errors and misconceptions in reasoning about Statistics and probability?' They assert that studies on errors and misconceptions related to statistical reasoning focus on topics associated with probability as well as misconceptions and errors related to additional topics such as contingency tables, sampling distributions and significance tests.

The problem is *more pronounced* in developing countries struggling with lack of resources as well as lack of innovation for making concepts such as those related to statistical inference easier for the students. Whereas in the developed world, there is now great emphasis on computer-based simulations for imparting these concepts, developing countries such as Pakistan are still very much into the theoretical way of delivering these concepts. Notwithstanding the fact that mathematical theory is absolutely important, there is no denying that software capable of generating simulations is *a wonderful tool* for elucidating mathematical results. In the absence of funds providing the capability to purchase expensive software for simulations, freely available/downloadable online software such as StatKey is 'just the right thing' for faculty-members and students of developing countries.

A few exceptions aside, learning from online webinars, etc. is still not a part of the 'academic culture' of the Pakistani society. Similar may be the case of a number of other developing countries. Faculty-members and students are happy to *get together* in order to learn from an expert who is right there, presenting in front of them, addressing them, answering their questions and putting questions to them. The workshop rendered by the author at ISOSS House, Lahore, Pakistan in November 2017 was meant to motivate faculty-members of various institutions to inculcate in their students' minds core concepts of statistical inference through *simulation-based procedures* using the freely available online software StatKey. Due to the positive experience and the highly encouraging feedback received from participants, the author is now motivated to *repeat* this workshop in a number of universities and higher education institutions located not only in Lahore but in various cities and towns of Pakistan. Exposure to this software may open to many '*a whole new world*' regarding a very interesting and highly effective way of learning and teaching some of the fundamental concepts of statistical inference.

REFERENCES

Batanero, C., Godino, J. D., Vallecillos, A., Green, D.R. and Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.

Garfield, J. & Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, 75(3), 372–396 doi:10.1111/j.1751-5823.2007.00029.x.

Habibullah, S.N. (2014). On the Commencement of a Culture of "Statistics Acceptance" in a Higher Education Institution Relatively New at Research. Invited paper in K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. iase-web.org [© 2014 ISI/IASE].

Morgan, K.L., Lock, R.H., Lock, P.F., Lock, E.F. and Lock, D.F. (2014). Statkey: Online Tools for Bootstrap Intervals and Randomization Tests. Invited paper in K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. iase-web.org [© 2014 ISI/IASE].

Onwuegbuzie A. J. (1997). Writing a Research Proposal: The Role of Library Anxiety, Statistics Anxiety and Composition Anxiety. *Library and Information Science Research*, 19(1), 5-33.

Suanpang, P., Petocz, P. & Kalceff, W. (2004). Student attitudes to learning business statistics: comparison of online and traditional methods. *Journal of Educational Technology & Society*, 7(3), 9–20.