ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 4, Issue 4, pp: (353-357), Month: October - December 2016, Available at: www.researchpublish.com

Teaching Statistics through Real Data: An Overview

Chuda Prasad Dhakal

PhD, Tribhuvan University, Institute of Agriculture and Animal Sciences, Rampur Campus, Chitwan, Nepal

Abstract: **On contrary to the yet most common practice so far, statistics teaching should not be taken as a mere fact of transferring information from one generation to another. Rather the goal should be, to habituate instructors and the students in statistical way of thinking for better statistics teaching/learning environment. Why statistics is learned and how could them be used in life sciences are the most fundamental factors to consider. This paper has an overview on how real data can play a significant role and why should them be considered in making statistics teaching more engaging, interesting and more meaningful to the students. Also, the impact of such teaching in the students' work situations have been discussed in brief. The paper also has discussed the benefits and the challenges while using real data in statistics teaching and as well have some in person experiences and the efforts explained while have come across struggling to make statistics teaching relatively satisfactory in a continuum.**

Keywords: **statistics teaching, real data, statistical way of thinking, data analysis, statistics education.**

I. INTRODUCTION

What makes a good statistics teaching / learning environment? Such questions would seldom get any candid answers. People commonly have different perspective answering such questions. So far, I have come across dozens of statistics teachers and thousands of students (I taught mathematics more than a decade in several high schools before I entered the university teaching service; TU, Institute of Forestry, Pokhara Campus 1998-2001; and TU, Institute of Agriculture and Animal Sciences, Chitwan, 2002-to the date). My experience with the nature of such obscurity, as well as the findings of [Grafield, 1995], [Garfield & Ahlgren, 1988], [Garfield et al., 2012], [Hannigan et al., 2013], [Neumann et al., 2013], [Grafield, 2016] suggests that statistics teaching in most of the cases could be made more interesting, more engaging and more contextual than how it has been till now. The recommendations in these papers in one or the other way are: instructors use real data in statistics teaching to emphasize concepts rather than computations. They focus more on applications and less on theory.

According to Guidelines for assessment and instruction in statistics education college report [2012], any teaching in statistics would make a 'statistical literary through the habit of making statistical thinking.' 'Conceptual understanding would be achieved via use of real data and analysing them with the available technology. Together with this, active learning in the class room and use of assessments to evaluate and improve the studies were also considered as the essential factors.'

Up to my graduate studies (BSc: Maths Physics and Statistics) I even did not end up thinking about what would make a good teaching. In occasions when there were hurdles in my studies, I felt that this was up to me and with my own reason. Either I was not meant for the specified subject or I was incapable or inefficient. No matter how hard I worked to make the subject/s interesting, I always failed them. And the ultimate way out throughout the college years was merely force memorizing the information given by the teachers. Finally, I graduated.

Next were my university days. The one and only reason for me to choose the subject 'statistics' for my master's studies was, compared to maths and physics, I smelled statistics relatively closer to the real-life context. This was encouraging

and motivating, it raised interest in me and gave me hope that I would pass the exam with ease. With their own stories and context, the university days passed. Same procedure in teaching was followed, readymade notes and methods and mode of instruction were replicated as in the bachelor's degree studies. My approach was the same, i.e. simple catching up of the information fed by the teachers without getting much knowledge in their application. However, I earned my Master's Degree in statistics in good grades.

My next move was clear. I worked hard and passed the very competitive Tribhuwan University (TU) Service Commission examinations and entered the teaching job in TU's, one of the institutes of applied sciences at the Institute of Forestry. Clearly, forestry students were aimed to pursue their studies in forestry, environment, ecology and so on, but not in statistics. I explored how statistics courses were and how the classes were taught so far. But, very frustratingly, I found them the same. Like past university statistics courses they were overly rigid and abstract. Teaching methods were not stimulating and enjoyable to make learning more effective. Conclusively, traditional methods of instruction were followed throughout!

However, by that time my eyes were open and I was totally convinced with the possibility of better methods for statistics teaching. I believed that statistics at any cost should be taught with learning process making more interesting, appealing and applicable. My big point was, statistics to the students of applied science should be a tool to deal with the substances they were related to. The students in forestry were meant for making their career in forestry, not for memorizing and reproducing the already established statistical theoretical relations and the formulas.

To get continued the story, any good teacher would be able to solve the exercises given in the text and prove the theorems. But, the poor teaching methods were neither meant to any application of the subject matter nor they took care to solve any real-life problems that the students would face at the start of their job. My intention was to make a breakthrough in this. The distinct shift I made was teaching through examples which had real data. I encouraged students to understand the root cause and the relationship between the variables. In the exam, I did not ask proofing of the algebraic expressions type questions. No any of them were to rote memorize. Instead, students wrote or explained the relationship between the given variables in non-technical language. Challengingly, in this period of my carrier building in teaching, I have faced across two interesting incidents. One was students boycotted my class in the name that I was unable to proofing identities, theorems etc. And the second was, they came out of the exam hall and burnt my effigy. The reason was, the questions I asked in the exam, were totally out of their expectations.

I was convinced in my way of teaching. It was vital for the students to solve real problems in their forthcoming work situations. Gradually my students and the college authority were put up in the issue. I explained them the benefits and the importance of the teaching methods I was following. Even though any formal evaluation of this non-conventional method of teaching statistics has not been carried so far, informal evidences are; the students now have self-realized the difference between these two methods of statistics teaching and those who have exposed to the later approach of teaching, have shown better performance in their work.

From this end, I have been continuously putting my effort to make statistics teaching meaningful at least in the case of applied science studies. I have been arguing over the fact that statistics teaching made more meaningful and substantial knowledge, and not only plain memorizing of unexciting facts and figures. My point there a great change in the teaching style was needed. But how? With what approaches? One significant factor to consider was teaching statistics through real data.

Both earlier and recent literature support this idea. Willet and Singer [1988] revealed that, more specifically, applied statistics courses are to teach how to address questions such as how to read other's research on related topics but not to become statisticians. Students in these cases are less interested in the algebraic ins-and-outs of mathematical statistics than they are in learning how to use statistics. Concerned educators should provide courses that meet the students' interests. In agreement to this idea, David et al [2012] has indicated several approaches to improve statistics teaching. They include more integration among course content, pedagogy, and technology. Their work also revealed specific teaching strategies to achieve goals in statistics teaching. These are computer-based activities and interactive multimedia. The importance of stimulating interest and determination in students was also noticed. These approaches emphasize on statistical thinking, using less theory and more data, making data analysis central, building intuition, fostering active learning, and using context to develop statistical inference.

According to [Hein, 1991] constructive theory of learning refers to the idea that learners construct knowledge for themselves. Each learner individually (and socially) constructs meaning as he or she learns. Constructing meaning is learning; there is no other kind. Two consequences of this view are: The teacher or the instructor should focus on the learner in thinking about learning (not on the subject/lesson to be taught) and presumed that there is no knowledge independent of the meaning attributed to experience (constructed) by the learner, or community of learners.

Tuning on this [Neumann et al., 2010] has had followings:

From a theoretical perspective, the application of real data in teaching statistics aligns several theories of learning. Within a constructivist theory of learning, students will construct knowledge based on their experiences using real data sets. New knowledge is integrated with previous knowledge regarding the interpretation of the data using relevant statistics. The data sets should be used so that the students have the opportunity to reflect upon their work with the data, with the teacher providing just enough guidance to assist the students to build their own understanding. Real data sets also provide a context to a statistical problem. Context may be based on data or it may be on the physical and social learning environment. Students may develop their statistical reasoning through an interaction between their contextual knowledge about the data set and their emerging statistical knowledge.

These literatures therefore, have shown the importance of application of real data in teaching statistics has been increasingly recommended in the field of statistics education. This has been proofed to have paramount benefits over traditional methods of teaching statistics. Real data sets can be used by students to practice calculations, gain experience in the interpretation of results, and develop their statistical reasoning about a problem. Teachers can use them to illustrate different research approaches, methods of data analysis, and applications of statistical theory to solve real-life problems.

II. BENEFITS

In the early years when no statistical packages or high speed computers were available, simple artificial data set were used to ease the computational burden. This approach of considering artificial data made the subject dull and unrelated to the students' practical interest. This induced them to believe that concern for numerical accuracy was of more importance than gaining conceptual understanding in the subject matter. According to Nick [2012] real data set related to appropriate topics can stimulate the students' interest on the practical use of statistics. It can be exciting! Students can then start reasoning on why any result might have happened and develop a habit of questioning the quality of data. By learning that reaching quality data is not a simple and easy process, they became better prepared for real data application in the real world.

Accordingly, this approach of teaching stimulates students to state research questions, select appropriate statistical models and interpretation of parameters estimates, interpret results, and analyse implications for policy and practices. Contextual data spark interest and allows students to assume the role of researcher by exploring data that address real research questions. Also, real data provide a practical arena in which students can learn how to link research questions to statistical models and how statistical methods can inform the current research debate. By using real data, we can teach not only how to analyse data, but also why to do so. Through real data we can teach applied statistics in the way that statistics are applied and these allow instructors to "model" good data -analytic practice, making statistics more pleasant to the students.

III. MOTIVATION

Students cannot experience the thrill of doing research when they analyse artificial data. They are not challenged to express their results in non-technical terms. All they do is perpetuate the myth that statistics is dry and dull. We therefore recommend that instructors and the textbook authors use real-world data so that students can learn skills in a realistic and relevant context. In this line, according to Willet and Singer [1998], one approach to make statistics teaching more interesting, more meaningful and more successful is, to capitalize on students' fascination, not for statistics itself, but for substantive problems that statistics can address. Meaning that, while thinking about teaching, teachers put themselves into the students' real level of thinking to accommodate a realistic method of teaching.

Real data sets can be obtained from students themselves. Students' generated data may have several benefits. Primarily this is likely to create a strong interest to the students. This engages students in developing their own research design, which increases interest in learning statistics techniques that help gain meaning from the data. However, care should be taken that, acquiring data in this way may not always be appropriate and it requires the students and teachers to invest

adequate time to achieve quality data. In this process the instructors and the students should be aware that best data sets: i) come in raw form, ii) are authentic, iii) include background information, iv) have case-identifying information, v) are intrinsically interesting or relevant, vi) are topical or controversial, vii) offer substantive learning, and viii) lend themselves to a variety of statistical analyses.

Neumann et al [2012] revealed that many teachers of statistics recommended "real data" during class lessons. The authors carried out research on the factors influencing the following aspects of using real data: a) relevance, b) interest, c) learning/remembering, d) motivation, e) involvement/engagement, and f) understanding of statistics. The research result indicated that both cognitive and affective factors were associated with using real data to teach statistics. In addition to this, the authors also have suggested the varied features in data sets, statistics teachers should look for, when designing their lessons.

IV. CHALLENGES

Using real data in teaching statistics is a challenging task. An instructor, in fact, will need varied but instant data sets relevant to different situation and context. A major persuasive factor for using artificial data is therefore that an instructor can readily create data sets with the requisite characteristics. On contrary to this, real data sets might be tedious to work on. For instance, they apparently increase the work load to the instructors, i.e. class preparation time, etc. Also, many a times real-life data sets, may be vital for students who have not had experience in their application. Real data can come in different form and need to be meticulously attained. Moreover, the sources of real data are to be carefully identified and explored. Obtaining real data in many occasions is an investigative task which is highly challenging.

V. CONCLUSION

In this paper, we reviewed that statistical concepts are best learned in the context of real data sets. Any data should be motivating and of interest to the students, showing how statistical techniques can uncover meaningful information from numbers. Through this discussion we conveyed a message that every instructor or any student at higher level should use real data for statistics teaching/learning activities. When real-life data is used to illustrate statistical notions, practice calculations, and show the application of statistics, there is some association with specific types of student learning experiences such as relevance, understanding and learning, interest, motivation and engagement. On contrary to the made-up data which reinforces the perception that statistics is artificial, lifeless, and uninteresting, the use of real data is a motivating tool that makes learning meaningful and prepares students to use statistical techniques in the real world.

Also, we shed light on that many authors have strongly argued the advantages of real data sets, approving, this can be a meaningful and effective vehicle for teaching statistics, enabling students to develop analytical skills through realistic research situation. Real-world data and an experimental research paradigm can be a strongest means for motivating students to learn not just how to analyse data but also why. Real data not only assists teachers in communicating how data is analysed but also why it is analysed. Briefly speaking, real data is argued to play a central role in contemporary teaching approaches that aim to develop statistical knowledge and reasoning skills in students.

The use of real data is challenging; however, the advantages far outweigh the difficulties of applying a teaching approach based on real data application. It is therefore, recommended that statistics teachers incorporate real data to make statistics teaching meaningful. Perhaps, the only way to experience the difference and to uncover the advantages of authentic data sets is to try them in the classes.

REFERENCES

[1] Garfield, J., & Ahlgren, A. (1998). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education,* 19 (1), pp. 44-63. doi: 10.2307/749110

[2] Garfield, J., delMas, R. & Zieffler, A. (2012). The International Journal on Mathematics Education 44 (7), pp. 883-895. doi:10.1007/s11858-012-0447-5

[3] Grafield, J. (1995). How students learn statistics. *International Statistical Review*, 63 (1), pp. 25-34. doi: 10.2307/1403775

[4] Grafield, J. (2016). Preparing to teach statistics. *ColleeBoard: AP Central. Retrieved from* http://apcentral. collegeboard.com/apc/members/courses/teachers_corner/28254.html Accessed on 05.09.2016

[5] Guidelines for assessment and instruction in statistics education, college report. (2012). American Statistical Association, USA.Retrieved from http://www.amstat.org/education/gaise/ Accessed on 17.08.2016

[6] Hannigan, A., Gill, O. & Leavy, A.M. (2013). An investigation of prospective secondary mathematics teachers' conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teachers Education*, 16 (6), pp. 427-449. doi:10.1007/s10857-013-9246-3

[7] Hein, E. G. (1991). Constructivist learning theory. *Institute for Inquiry; Lesley College. Massachusetts USA*. Retrieved from http://www.exploratorium.edu/education/ifi/constructivist-learning Accessed on 19.10.2016

[8] [8].Neumann, D. L., Hood, M., & Neumann, M. M. (2010.). Using real-life data when teaching statistics: Student Perceptions of this strategy in an introductory statistics course. *Statistics Education Research Journal*, 12(2), pp. 59-70. Retrieved from http://iase web.org/Publicationsphp?p=SERJ Accessed on 17.08.2016

[9] Nic. (2012). Stop faking it! Data should be real. *Statistics Learning Centre: Learn and Teach Statistics and Operations Research*. Retrieved from https://learnandteachstatistics.wordpress.com/2012/04/23/stop-faking-it/ Accessed on 17.08.2016

[10] Willett, J. B., & Singer, J.D. (1988). Providing a Statistical "Model": Teaching Applied Statistics using Real-World Data. *Harvard University Graduate School of Education, Mathematical Association of America, MAA Note* (26), pp. 83-98. Retrieved from gseacademic.harvard.edu/…/Willett Singer… Accessed on 17.08.2016