"MINI-STUDIES OF PREVALENCE": A WAY TO INTRODUCE PROBABILITY BASES AND THEIR RELATIONSHIP WITH DATA ANALYSIS

Emilse Gomez-Torres
Statistics Department, Universidad Nacional de Colombia
egomezt@unal.edu.co

This paper exposes the design of a formative activity, and the students' achievement, in an introductory course to biostatistics offered in 2021 at a Colombian public university, under the modality of remote classes (due to pandemic restrictions). The activity intended to promote probabilistic reasoning and to develop some fundamental statistical skills, required for the future professional practices of the learners. The task was related to the estimation of the prevalence of a characteristic associated with physical activity habits or sedentary behavior for a subpopulation of teenagers, delimited for academic purposes. The achievement of learning objectives differs between groups of students; the teamwork seems to be related with performance differences, as well as the self-perception of some study habits' favorability.

INTRODUCTION

The prevalence of a disease is a common concept on the health area (human or animal), associated with the probability that an individual (person or animal) belonging to a certain population (in a bounded spatiotemporal context) presents such disease; therefore, prevalence studies make it easier to showing the connection between descriptive statistics and probability. In fact, the prevalence of a disease had been suggested as practical application to teaching conditional probability, giving to students a statement and a table of contingencies (for example Sahai and Reesal, 1992).

However, in professional practice, someone rarely has access to the complete population; for that reason, each prevalence of interest is usually approximated by analyzing data from sampling; thus, prevalence studies help to the introduction of estimation in connection with probability. With this vision, within a course of introduction to biostatistics, the students were asked to carry out a mini-study of prevalence for two or three characteristics contained in a Colombian study on physical activity and sedentary behavior in teenagers, whose database was freely accessible and was edited by the teacher for training purposes.

This formative activity uses real data with context and purpose, aligned to scholar level, as suggested GAISE (2016) to take advantage of the increase of available open databases. Also, the activity was planned for team work and following flipped classroom scheme. In the scenario of remote classes, it arose as a question: Did the designed task promote in students probabilistic reasoning and fundamental statistical skills useful in their future professional practice? Consequently, the aim of this research was to analyze the performance of the participants in the implementation of this activity.

Below are synthetically exposed some proposals on training in Biostatistics for undergraduate students and experiences with flipped classroom. Then, the methodology followed in this classroom research is presented, after that the observed results are exposed, and it ends with a discussion about the scope of this research.

BACKGROUND

Schwartz et al. (2016) summarized experiences shared in 2014 during the panel *Flipping the Biostatistics Classroom.* The authors acknowledge that, from the research point of view, it is necessary to deepen the effects of the flipped classroom on student learning and performance.

Fuentes et al. (2020) explored the effect of various non-academic aspects on the learning of 360 university students (18-23 years old) who participated in courses under the flipped classroom scheme. According to the evidence, the effectiveness of learning in the flipped classroom has a negative association with the workload of the students and a positive association with the family context, autonomy in daily activities, age, self-esteem and motivation, in this order of influence.

METHODS

Mixed research methodology with descriptive analysis was applied. At the beginning of the academic period, scholar information was collected from all students enrolled in the course; to identify

In: EM Jones (Ed.), *Fostering Learning of Statistics and Data Science*
Proceedings of the Satellite conference of the International Association for Statistical Education (IASE),
July 2023, Toronto, Canada. https://iase-web.org/

patterns in these data were used quantitative methods. After the middle of the academic period, the reports presented by teams were valued quantitatively and qualitatively with an evaluation matrix.

Instruments

Two collection instruments were used: the record of the academic management of the course and an online questionnaire. The first informed about student performance, particularly grades for each criterion of the evaluation matrix (with their feedback), the final grade, and the level of class participation. For the qualitative analysis, the written productions of the teams were used.

The online questionnaire was designed by the teacher-researcher to characterize the students enrolled at the beginning of the course. This instrument was a structured questionnaire with 41 questions (58 variables), 33 closed and 8 open, distributed in four sections: academic or personal information (17 variables), previous knowledge (14 variables), study habits in the remote class setting (9 variables), perception of pedagogical activities' favorability in the remote class setting (18 variables)..

Participants

The formative activity was implemented in a course of introduction to biostatistics during December of 2021. In the first week of classes (October 2021), 51 people were on the list, 49 of whom answered the online questionnaire. Their self-perception about proportional reasoning was relatively high, 6.1% classified themselves at a low level, 14.3% at a medium level, 44.9% at a medium-high level, and 34.7% at a high level. Their self-perception of handling spreadsheets was more diverse: 12.2% classified themselves at a low level, 36.7% at a medium level, 28.6% at a medium-high level, and 22.5% at a high level.

The prior knowledge of the students was very diverse at the beginning. Regarding training in mathematics: 63.3% had studied basic mathematics at the university level, 40.8% differential calculus, 6.1% integral calculus, and 26.5% only had school-level mathematics. Regarding training in statistics at the school level: 14.3% attended a specific subject, 42.9% recognized only the content included in the mathematics subject, 40.8% expressed zero contact with statistics at the school; on the other hand, 59.2% had attended a course of general statistics at the university level; in summary, 6.1% of the group had not had prior contact with statistics.

For this study, the observed sample is the set of 26 people who submitted the mini-study of prevalence report and had completed the initial questionnaire.

Analysis procedures

Responses to open questions from the initial questionnaire followed a coding process: each sentence was categorized with a short tag, then all tags were reviewed and classified to identify patterns, some of which generated macro categories, which were subsequently interpreted.

The macro-categories, the closed questions and the performance in the course were gathered in a single database to facilitate the joint descriptive analysis; statistical measures were used according to the type of variable and the small sample size.

DESIGN OF THE FORMATIVE ACTIVITY

The experience was designed in order to provide the probabilistic bases to understand concepts and procedures of data analysis required in prevalence studies, a statistical application common to the field of health (human or animal) that was considered motivating for first-year or second-year students with various levels of prior knowledge and academic interests.

The activity was proposed to be developed in work teams during two sessions of flipped classroom, following the jigsaw strategy, and was called a "mini-study of prevalence", whose aim was to study (in an introductory way) the prevalence of two or three attributes in a target population; that would be presented in an article-type report of a maximum of three pages. The guides for the mini-study and the matrix for evaluating the reports were released some weeks before the practice.

The target population of the mini-study (subpopulation of Colombian teenagers) was assigned to each student. The attributes to analyze their prevalence had to be selected by each work team from the variables available in a file provided by the teacher-researcher. This file was an edited extract from the database on physical activity and sedentary behaviors of teenagers, one of the chapters in ENSIN

2015 (acronym for *Encuesta Nacional de la Situación Nutricional*, that means National Survey of the Nutritional Situation, carried out in Colombia during the year 2015). The file had 18 of the 107 variables of original database.

According to the jigsaw, the individual contribution was linked to a different task for each member; in this case, the database was divided into six spreadsheets (92-114 rows), one for each geopolitical region of the country, and each student was assigned one region. Each student selected two attributes of interest, compute as necessary to estimate its prevalence (punctual and by interval).

The work teams had 3 or 4 members, the region would not be repeated and together they would analyze the same attributes for various regions of the country. The team had to decide which two or three attributes would include in its report according to the relevance or the implications in the context; as each member previously select ones, not necessarily the same, the team had to discuss about it and, when the decision was made, students had to adjust their input to the work in order to get a unified report. After that, among members had to reflect about results, variability and sources; the team analysis looked for the recognition of variability between regions and possible sources of variation; the individual analysis promoted the recognition of patterns and possible sources of variation within each region.

The guides given by the teacher-researcher suggested a sequence of steps to follow for the preparation of the individual contribution (first session of the practice week), as well as a sequence of steps to carry out the group activity (second session of that week). During each session, the teacher-researcher asked if there were doubts and tried to solve them, to the whole class or to each team.

The evaluation matrix for the report consisted of nine criteria. Four referred to aspects of form (20% of the grade) linked to the development of soft skills. The other five criteria valued the statistical application in the context of estimating a prevalence (80% of the grade).

RESULTS

Description of the participating sample

The 26 participants were classified into four subgroups based on their performance in the course and the stability of their work team. Table 1 presents summary measures for quantitative variables of interest for the entire group of 26 participants and broken down by subgroup.

- SG1 was formed by biology, physiotherapy, speech therapy and veterinary medicine students, who obtained a grade higher than 4 in the mini-study report, approved the course and their work team was kept until the end of the academic period. It is the second most variable group in terms of number of enrollments and percentage of progress in the undergraduate program.
- SG2 was formed by pharmacy and nutrition and dietetics students, who obtained a grade above 3 in the mini-study report, approved the course, and their work team changed throughout the academic period. It is the smallest and most homogeneous group in terms of number of enrollments and percentage of progress in the undergraduate program.

Table 1. Description of the participants globally and by subgroups through variables related to academic performance

Descriptive statistic	SG1	SG2	SG3	SG4	Global
Number of participants	9	3	7	7	26
Average Final Note (s.d.)	4.26 (0.45)	3.96 (0.48)	3.62 (0.31)	2.14 (0.41)	3.48 (0.95)
Average Grade of the Report (s.d.)	4.34 (0.20)	3.86 (0.32)	3.42 (0.20)	3.86 (0.35)	3.91 (0.45)
Number of registrations (Min – Max)	2 - 7	2 - 4	2 - 9	1 - 4	1 - 9
% of undergraduate progress (Min – Max)	9.2 - 47.6	8.2 - 22	6.9 – 75	0 - 16.1	0 - 75

- SG3 was formed by students of biology, physiotherapy, speech therapy, geography, chemical engineering, veterinary medicine and chemistry, who obtained a grade between 3 and 4 in the mini-study report, approved the course and their work team was kept until the end of the academic period. It is the most variable group in terms of number of enrollments and percentage of progress in the undergraduate program.

- SG4 was formed by biology, pharmacy, speech therapy, and chemistry students who obtained a grade higher than 3 in the mini-study report and dropped out of the course after middle of the academic period. It is a relatively homogeneous group in terms of their time at the University, with the students with the least progress in the undergraduate program.

The time available for the course was measured through three variables: work dedication, weekly hours planned for the course and the number of credits enrolled to the academic period (in Colombia, one credit is equivalent to 48 hours of academic work, adding class time and autonomous work time). Table 2 shows that, at the beginning, they weekly allocated around 8.3 hours on average for biostatistics (theoretically, should be 9 hours); also, the number of enrolled credits on the academic period is very heterogeneous, more diverse in SG1 and SG4 than in the other two subgroups.

Table 2. Description of the participants globally and by subgroups through variables related to the time available for the course

Descriptive statistic	SG1	SG2	SG3	SG4	Global
Number of participants	9	3	7	7	26
Average of enrolled credits	15.2 (1.87)	12.0 (1.41)	16.6 (0.73)	13.6 (2.61)	14.8 (2.37)
Average of planned weekly hours	8.22 (0.79)	8.33 (0.47)	8.14 (0.64)	8.57 (1.40)	8.31 (0.95)

In this sense, the really available time for the course is very diverse among the participants: Twelve of the 26 participants stated that they were working, two full-time (both are part of SG4) and ten part-time. The two people who worked full time enrolled to this academic period 11 or 12 credits (that means, respectively, 33 and 36 hours per week to study) and stated that they expected to having 8 or 10 hours for the biostatistics course. Two students expressed working part-time and enrolled 18 and 19 credits (that means respectively 54 and 57 hours per week to study), which seems like an excessive burden.

Finally, the reasons to select this specific class (among 14 of introduction to biostatistics with other schedules and teachers) were classified in macro categories of arguments. These showed high diversity in SG1 and SG3, referring to academic, curricular, logistical, and operational issues. In SG4 operational reasons predominated (5 out of 7), for example "learning to use the different platforms of the subject to have a good management of them in my daily and professional life" (response from E33). In SG2, academic (2 of 3) and operational (1 of 3) arguments were observed.

Participant performance

The grades of the reports delivered by the students showed high homogeneity in five of the nine criteria: presentation of the research questions made by the work team; application of statistical procedures; analysis of variability between regions and within each region; appropriate use of statistical and probabilistic language; compliance of grammar and writing rules.

It was easier to recognize variability among regions than within each region. The reflections about sources of variation showed their knowledge of context, for example:

> The difference between the probability of attending school between the Eastern region and the rest of the regions is notable, with this having the lowest probability (0.489)[…] This particular behavior of the region can be associated with the differences that it presents with Bogota and the central region, due to these two regions has a high level of urbanization, where sedentary behavior tends, while the eastern region is a more rural region where sedentary tendencies are less. (Ford team)

It is considered that these 26 students achieved most of the learning objectives set for the activity. Notice that in the most of reports, the procedures were showed using the same scheme and table used by the teacher-researcher during the class in which estimation was explained and the procedural issues were illustrated with other region and other variable of the same survey, but not included in their database.

Regarding the criterion linked to the appropriate use of statistical and probabilistic language, in general the performance was good, the most frequent failures were related to the use of colloquial terms and a writing that tended to describe without mention of uncertainty; for example: “the number of people who see themselves as fat is greater than the number of people who try to maintain their

body weight. There is a causality between trying to maintain body weight and seeing yourself as fat." (Crick team)

The four criteria that made the difference in the grades were: compliance with instructions; coherence in the document; description of the method followed to answer the research questions; assessment of the reliability of the results and its effect in the conclusions. The performance of the teams that obtained a low score in these criteria was affected by deliver very late; each wrote a part; omission of information or assessment, or absence of arguments for decisions made.

The reflections about reliability of the results and its implications were focus on recognition of limitations of their mini-study: "we do not know if, within the different regions on which we base our study, factors such as the climate, the regional economy or the pad between the home and the school, considerably influence in the answers that teenagers give in the survey." (Snedecor team)

Some students recognized the effect of sample size in the confidence interval, for example:

> 33.3% [94% CI 18.6-48.1] of teenagers who while enrolled spend more than 3 hours a day playing video games or in front of a digital screen in activities not related to work academic environment during a school week.[…] surprising that only 31.6% of surveyed are enrolled in an educational institution, this compromises the estimates made to teenagers who are enrolled and perform physical activity together with teenagers who spend more than 3 hours in front of screens, since there is such a small sample, the error sampling increases considerably compromising the confidence interval. (Fisher team)

The last paragraph contrast with the next, that showed a deterministic vision and overgeneralization.

> For the Bogotá region, […] which means that there is a non-negligible percentage that performs less exercise for less than 3 days and in some cases does not perform any physical activity, this comes hand in hand with a percentage of 64% who spend more than 3 hours in front of a digital screen. The percentages are worrying in terms of physical activity, there is evidence of a sedentary lifestyle in teenagers because they spend many hours in front of a digital screen. (Fisher team)

Both pieces were taken from the same team, but showing different level of knowledge, were an example of low coherence on the report due to distribution of tasks without joint revision. This team worked only collaboratively, something went wrong to achieve cooperation.

Relationships between prior information and performance in the mini-study of prevalence report

The following analyzes are based on a reclassification of the 26 participants considering only their grade in the mini-study report, grade between 4 and 5 means good performance, and grade between 3 and 4 corresponds to acceptable performance.

The previous contact with statistics tends to be a little higher among the participants with good performance (5 out of 14 had received statistical instruction in their school and university training), while among the participants with acceptable performance predominated the completion of only one course at the university level (5 of 12).

Mechanisms for taking notes denote some differences. Among the participants with good performance, almost all expressed using a notebook (12 of 14); while among the participants with the acceptable performance, half also expressed using notebook (6 out of 12), one expressed not taking notes and the others were distributed among other options.

The assessment of software exploration as part of the autonomous work in previous courses showed differences; it tends to be more valued among the participants with acceptable performance (with a predominance of the favorable rating, 6 of 12) than among good performers, whose answers were more diverse, including one person who considered it unfavorable. The evaluation of guide readings given by the professors of previous courses tends to be favorable (8 out of 14) among the participants with good performance; it was distributed almost equal among unfavorable (5 of 11) and very favorable (4 of 11) among the acceptable performers.

The assessment of explanation recorded by teachers from previous courses was distributed almost equally from unfavorable to very favorable among good performing participants, while the acceptable performers rated it equal as favorable and very favorable (5 out of 10, each). The evaluation of asynchronous class in previous courses tended to be very favorable (8 of 12) among the acceptable performers, while it was distributed almost evenly from unfavorable to very favorable among good performers.

The positive perception of the high favorability of software exploration as part of the autonomous work, the recorded explanation and the asynchronous class among the participants with

acceptable performance, as well as their low preference for reading guides, may be related to their handling study time. The students informally expressed that they liked having recordings because they could advance it or modify the speed in order to "go" to the part of the class in which the indications related to the development of the work were specifically explained.

It should be noted that just under 20% of the participants omitted to assess activities related to discussion with peers as part of their teamwork. These aspects were considered relevant for this research due to the mini-study of prevalence report was posed as the product of cooperative learning through jigsaw strategy; previous experiences of team work or collaborative work could affect the study habits developed before facing the task proposed in this course.

DISCUSSION

This activity was designed in the context of a Colombian public university, with a high diversity of academics, demographics, and professional interests, in order to promote students' probabilistic reasoning and fundamental statistical skills useful in their future professional practice. The analysis of the delivered reports by participants' teams showed different levels of achievement for some learning objectives. Also, based on an exploratory analysis of academic variables self-reported by learners, the teamwork and the perception of some study habits' favorability, under remote class conditions, could be related with performance differences.

The flipped classroom modality with team work promotes cooperative work and a differentiated gain is observed between students (Schwartz et al, 2016): participants with lower bases or less developed skills learn from their peers due to easier and faster interaction with them than with the teacher, while the most advantaged participants deepen their knowledge by explaining a particular topic. However, in this classroom experience, teamwork was assumed in different ways among the students. The productions of the participants suggest that, in SG1, each team articulated to achieve cooperative learning, in accordance with the design of the task. Meanwhile, in the other subgroups, it seems that several people were associated without a team dynamic, whose only common purpose was to present a report. Highlight that cooperative learning requires coordinating, planning and interacting, while collaborative work can omit the last two actions.

Besides, the quality of the mini-study reports could be reflecting the commitment of the participants with their learning process or their available time for extra-class tasks; as Fuentes et al (2020) concluded, the student's workload negatively affected the effectiveness of learning under the flipped classroom scheme.

One limitation of this study is the low number of participants, who made up a convenience sample. Another limitation depends on the unusual circumstances of the implementation, due to the pandemic; in particular, on implementation days, many students did not connect to class, leaving doubts about how their teams carried out teamwork.

REFERENCES

Fuentes, A., López, J., Parra, M.E. & Morales, M.B. (2020). Diseño, validación y aplicación de un cuestionario para medir la influencia de factores exógenos sobre la eficacia del aprendizaje invertido. *Psychology, Society, & Education*, *12*(1), 1-17. https://doi.org/10.25115/psye.v0i0.2334

GAISE. (2016). *Guidelines for Assessment and Instruction in Statistics Education College Report*, Tech. Rep., ASA Revision Committee. Available on http://www.amstat.org/education/gaise.

Sahai, H., & Reesal, M. R. (1992). Teaching elementary probability and statistics: some applications in epidemiology. *School Science and Mathematics*, *92*(3), 145.

Schwartz, T., Andridge, R., Sainani, K., Stangle, D., & Neely, M. (2016). Diverse Perspectives on a Flipped Biostatistics Classroom. *Journal of Statistics Education*, *24*(2), 74-84. https://doi.org/10.1080/10691898.2016.1192362