HIGH SCHOOL TEACHERS´ STATISTICAL REASONING ABOUT COMPARISON OF DISTRIBUTIONS OF DATA IN A COMPUTER ENVIRONMENT

Santiago Inzunsa and José A. Juárez
Autonomous University of Sinaloa, México
sinzunza@uas.edu.mx

The article presents results of a study about statistical reasoning developed by seven high school Mexican teachers on how teachers compare distributions of data in a computer environment. The evaluation instruments adopted were four activities carefully chosen as well interviews with two of the teachers involved. The results indicate that the statistical reasoning of the teachers was predominantly oriented to inferior levels of the model SOLO (Structured Observed Learning Outcomes); in particular the process of informal inferences was the one that produced more difficulties for them. The teachers did better on working with tables and averages in comparison to the distributions. Perhaps the reason why is that they didn't make the most of the software potential for the manipulation of several graphic representations that would allow a better and more complete analysis.

INTRODUCTION

The comparison of distributions is a theme which provides contexts where the students can use several concepts (e.g. central tendency, distribution, variability) and statistical reasoning elements in order to respond to significant questions related to the data. Comparing distributions of data is of such an importance in statistics that a hypothesis testing –one of the principal inference methods-, consists in essence of comparing two distributions: the distribution obtained from a dataset gathered from a sample and an assumed distribution taken from a population. Makar & Confrey (2004) assert that the comparison of two distributions requires focusing both on the central tendency and on the distributions. They go on to say that this process provides ideas about some other statistical concepts and that it helps develop several aspects of statistical reasoning. Konold & Higgins (2002) argue that tasks focused on the comparison of data can help students to start looking at data as a distribution, instead of seeing them only as isolated elements. In addition, data comparison tasks can make students take into consideration not only central tendency measures, but also variability as part of its description. In this context, and considering that computer technology is becoming increasingly important in teaching statistics, this research study has set out to investigate the levels of statistical reasoning developed by high school teachers as they compare distributions of data in a computer environment as Fathom software (Finzer et al., 2002). In particular, we address the following questions: What aspects of the distributions do teachers consider to make its descriptions and comparisons? Which elements of an informal statistical inference do teachers employ to justify their conclusions? And, at what level do teachers utilize technology to represent, organize and reduce datasets?

THEORETICAL FRAMEWORK

Distribution refers to the arrangement of values of a variable along a scale of measurement resulting in a representation of the observed or theoretical frequency of an event. Descriptive statistics are indices of distribution, they summarize complex data into measures that can be compared against each other to ascertain the nature of a dataset, and the degree to which two or more datasets are similar (Leavy, 2006). Graphical representations serve as useful tools to communicate aspects of a distribution as they facilitate a focus on aspects of the data that may be missed with the use of descriptive statistics alone. The concepts and processes aforementioned may be used by the subjects in different levels of complexity, based on the comprehension of the process of comparison of distributions. A taxonomic model that allows us to account for it, and has been used to establish categories of cognitive development on various statistical concepts, is the SOLO model (Structure of Observed Learning Outcomes). This model, which was developed by Biggs and Collis (1982), involves five levels: prestructural, unistructural, multistructural, relational and extended abstract, through which one can assess the sophistication and complexity with which subjects combine the various concepts and processes required in comparing distributions. This

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

taxonomy is a model which is based on the idea that in the progress from incompetent to expert, subjects exhibit a consistent sequence or cycle of learning. This sequence refers to an increase in the structural complexity hierarchy of responses to a given task. For comparison of distributions we made a summary of the categories that some authors (e.g. Langrall & Mooney, 2002; Bakker & Gravemeijer, 2004) have defined for several of the concepts involved, and we have adapted the use of them in a computer environment (see Table 1).

Table 1: Model SOLO levels applied to the comparison of distributions of data in a computer environment.

Phase	*Representation/Organization/Reduction*	
Data Analysis	*Level*	*Characteristics*
	Prestructural (P)	Calculates statistics and creates the graphics the software performs by default, which may not be adequate.
	Unistructural (U)	Construct a graph but does not try to rearrange the data for better graphical representation for other conclusions. Use the statistics obtained by default.
	Multistructural (M)	Builds two or more graphic but do not use them for additional information allowing new interpretations. Seeks to calculate additional statistics.
	Relational (R)	Significantly manages multiple representations, graphs and tables, organizing and reorganizing the data.
	Description/Comparison	
	Level	*Characteristics*
	Prestructural (P)	Doesn't make reference to key aspects to formulate descriptions or comparisons of distributions.
	Unistructural (U)	Describes and compares the distributions focusing on one key aspect of reasoning (center, shape, gap/overlap, dispersion and individual cases).
	Multistructural (M)	Describes and compares the distributions over focusing on more than one key aspect of reasoning, but without integrating them.
	Relational (R)	Describes and compares the distributions establishing relational connections between several key reasoning aspects.
Conclusions	*Informal inferences*	
	Level	*Characteristics*
	Prestructural (P)	Doesn't pose inferences, and when it comes to posing them, it does not address key aspects of casual inference (generalization, data, statistical concepts, probabilistic language, sampling and explanations).
	Unistructural (U)	It raises a conclusion, it can refer to data and statistical concepts but not generalizations are made.
	Multistructural (M)	It raises a conclusion and makes an inferential statement (generalization) referring to two or more other key aspects, without achieving the integration of what was declared.
	Relational (R)	It poses a conclusion and makes an inferential statement (generalization) described in a way that it connects all aspects of a casual inference.

METHODOLOGY

The context of the study was a refresher course on exploratory data analysis which involved seven high school teachers of statistics. The evaluation instruments used were four activities carefully chosen as well as interviews with two of the teachers involved. For the analysis of reasoning developed by the teachers, the concepts and processes described in Table 1 were considered. Each of these constructs integrate various indicators (center, shape, gap, overlap, dispersion and individual cases) that help provide a detailed description of the reasoning. The following describes an activity used:

Cholesterol levels are a risk factor for heart disease. The data shown below are for cholesterol levels of 24 patients before and after taking a diet based on the consumption of vegetables for a month.

- *Analyze the data thoroughly. Use all the available resources you consider appropriate.*
- *Doctors recommend a cholesterol level below 200, what percentage of patients are above this value before and after the diet?*
- *Based on the analysis of the data, do you think that a vegetarian diet was effective in reducing cholesterol? Explain.*

RESULTS AND DISCUSSION

According to the data of Table 2, multistructural and unistructural levels in the use of representations for the organization and reduction of data as well as processes of comparison and description predominated. Further, this indicates that one or two graphs were used. In some cases additional statistics provided by the software by default were also used for the comparison of the distributions, but without integrating them to a more complete analysis of the data. The process that posed more difficulties for teachers was the informal inferences as they did not use generalizations to the populations from which the data come. This shows that they were located mainly at the unistructural level.

Table 2: Levels of SOLO model reached by each teacher in each of the constructs of the comparison of distributions

Teacher	*Representations Reduction/Organization*					*Description/Comparation*					*Informal Inferences*			
	Activities					Activities					Activities			
	1	2	3	4		1	2	3	4		1	2	3	4
1	M	M	M	M		M	M	M	R		R	P	U	U
2	U	U	U	M		M	U	U	M		U	U	U	U
3	M	M	M	M		U	M	M	M		U	U	U	U
4	M	M	U	U		M	M	M	M		U	P	U	U
5	U	M	U	U		M	M	U	M		U	U	U	U
6	M	M	U	U		M	M	U	M		U	U	U	U
7	U					M					M			

Meanwhile, based on the data in Table 3, it is assumed that several indicators for a proper comparison of distributions were not considered by the teachers. The Indicators other cutoffs and the arithmetic mean were the most used elements, followed by informal language about dispersion, standard deviation and minimum and maximum values.

Table 3: Frequency of use of indicators of the reasoning of the process Description-Comparison.

Activity	*Center*			*Shape*						*Gaps /Overlap*		*Dispersion*				*Individual cases*		
	C1	C2	C3	F1	F2	F3	F4	F5	F6	D	T	D1	D2	D3	D4	V1	V2	V3
1	6	3				1	7			1		1	1				3	
2	6	1					5						2		2	3	2	1
3	4			1			5					1	1		2		1	
4	4						5					2	4		6	1		

C1: Mean
C2: Median
C3: Mode

F1: Modal Grouping
F2: 50% Central
F3: Quartiles
F4: Different Cut-Off Points
F5: Bias
F6: Shape

D1: Range
D2: Standard Deviation
D3: Interquartile Range
D4: Informal Language

V1: Outliers
V2: Maximums & Minimums
V3: Other Values
D: Gaps
T: Overlapping

Some typical representations used recurrently teachers are the dot plots and statistical tables as shown below:

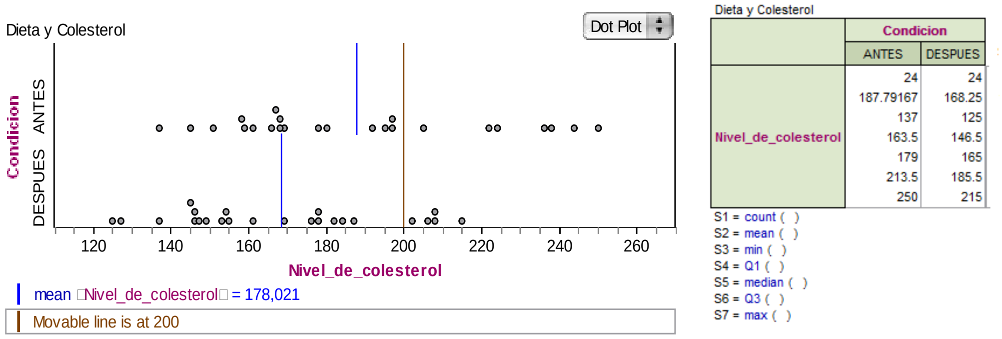

Fig. 1: Graphs and tables created by Teacher 1

For example, one teacher, who had the highest levels of reasoning, in the context of the cholesterol diet activity (see Figure 1) created a dot plot indicating the means of each distribution and the cutoff point of cholesterol equal to 200, which was supplemented by a table with five summary measures (count, mean, median, maximum and quartiles).

CONCLUSIONS

The results of the study show that teachers who participated in the study have little knowledge about the implicit nature of a description and comparison of distributions, since 18 indicators mainly used the average and the special cut-off points, followed by a moderate use of the standard deviation, range and outliers. As for casual inference indicators the teachers exhibited reasoning at a unistructural level; that is, many answers were at sampling level without attempting to raise an educated guess about what would happen to the population level. It was noted that teachers did not use the full cognitive potential that the software provides through the multiplicity of dynamic representations of the data as they frequently limited themselves to the construction of graphical and tabular representations with averages displayed by default by the software. According to the above mentioned, teachers need to go beyond the descriptive view of the data and adopt a more exploratory vision through the use of computational tools.

REFERENCES

Bakker, A., & Gravemeijer, K. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Finzer, W., Erickson, T., & Binker, J. (2002). *Fathom Dynamic Statistics Software*. Emerville CA. Key Curriculum Press Technologies.

Konold, C., & Higgins, T. (2002). Highlights of related research. In S. J. Russell, D. Schifter & V. Bastable (Eds.), *Developing mathematical ideas: Working with data* (pp. 165-201). Parsippany, NJ: Seymour Publications.

Langrall, C. W., & Mooney, E. S. (2002). The development of a framework characterizing middle school student´s statistical thinking. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Voorburg, the Netherlands: ISI.

Leavy, A. (2006). Using data comparison to support a focus on distribution: Examining preservice teachers´ understanding of distribution when engaged in statistical inquiry. *Statistics Education Research Journal*, *5*(2), 89-114. Retrieved from: http://www.stat.auckland.ac.nz/serj

Makar, K. M., & Confrey, J. (2004). Secondary teachers reasoning about comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking*. Dordrecht, the Netherlands: Kluwer Academic Publishers.