STATISTICS AND MATHEMATICS:
IS IT POSSIBLE TO CREATE FRUITFUL LINKS?

Linda Gattuso
Université du Québec à Montréal, Canada
gattuso.linda@uqam.ca

Statistics is generally taught in schools as part of the mathematical curriculum and, as a result, statistics is perceived as a mathematical concept. Moreover, the teaching emphasis is often placed on the computation of statistical information instead of the development of an " authentic data analysis point of view" (Cobb, 1999). In this paper we consider whether a constructive relationship between statistics and mathematics is possible? We examine examples of statistical teaching situations and analyze them both from a mathematical content point of view and from a statistical perspective, to point out that the interaction between mathematics and statistics is in fact feasible/(practicable). To achieve this, it is necessary to complete a mathematical conceptual analysis of statistical concepts. It is essential to highlight the mathematical concepts beneath the statistical concepts in order to link both of them in creating classroom activities.

INTRODUCTION

It is not necessary to emphasize the importance of statistical knowledge for today's citizen and we agree with Konold (2003) that without it, it is difficult to have an informed opinion and participate in social and political debates concerning environment, health, education, etc. On the other hand, many teachers are not comfortable with the subject. They lack confidence because they are not familiar with the mathematical content and data analysis (Russell, 1990a). Teachers that do not have a specific training get their information directly from school textbooks and may not detect the errors present (Hawkins *et al*., 1992, p. 286). Consequently, teachers are unable to recognise and respond to students, conceptions and misconceptions (Russell, ibid.)

Besides, the need to teach a subject not well mastered generates apprehension. Interviews of Italian mathematics teachers reveal "a feeling of insecurity due not so much to a lack of preparation in statistics, but also to a lack of preparation in the teaching of statistics (Gattuso an Pannonne, 2002). Additionally, teachers are not aware of the richness of the statistical content they have to teach. For example, descriptive statistics is seen as easy and not very interesting. Consequently, statistics is relegated at the end of the school year if not completely forgotten because there is no time (Aksu, 1990).

In this paper, we will try to call attention to the fact that statistics can also assist the learning of mathematics because it encases many mathematical concepts in a realistic and motivating context. In fact, much of mathematics is developed to describe and model phenomena of life, from day to day accounting to medicine and economy, and they are often included in statistics. It is necessary to match mathematical concepts to their applications in statistics so that one supports the development of the other (Dunkels 1990).

We also think that it is important to pay attention to the teachers uncertainties by assuring them that, while doing statistics, they are really doing a lot of mathematics. To do so, we will analyse some statistical teaching activities for elementary and secondary school and underline the mathematical concepts encompassed.

A FIRST STEP: SIMPLE DATA ANALYSIS

Initially, mathematics was developed to face everyday problems. Numbers are found in everyday life in comparisons, tables, orders of magnitude, rounding, estimating, prices and other numerical messages. So "It is important to develop children's mental images of numbers parallel to their acquisition of counting and calculation skills" (Dunkels, 1986, p. 61). Statistics that try to describe the world around us enclose much of these mathematics. An analysis of some school textbooks and descriptions of experiments reveal that statistics contain most of the mathematics included in the elementary curriculum and a part of the high-school curriculum. Data analysis usually begins with descriptive statistics, particularly with small surveys. These simple activities can be profitable even for preschoolers. Dunkels (1991) describes a "fruit survey experimented with 6 years old. At first children record the numbers of fruits of each kind they brought at school for their snack by drawing each fruit on a cardboard. In a second step, using a one to one correspondence, they begin a counting experience. The children handle numbers in a realistic context that favour the construction of the concept of number. After that, a variety of paths are possible.

One of them is symbolising. It can be as simple as replacing each "fruit" by a dot.

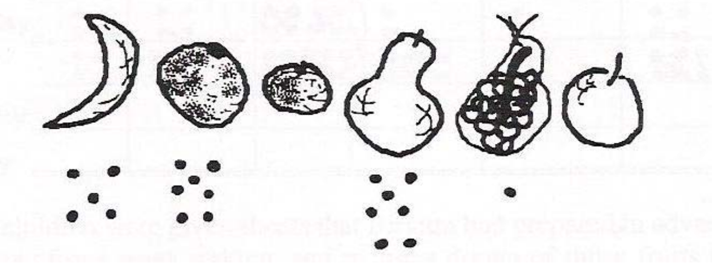

Figure 1: The fruit diagram (Dunkels, 1991)

Normally, 4-8 years old are not ready to pull their attention away from individuals (Russell, 1990), but collecting data and organising it systematically make them realise the existence of others and of differences. This can constitute an initiation to the concept of variability that is constantly active in data handling activities.

The construction of graphical representations is also fertile. In a first step, it is important to let the children create their own representations, and from there bring them to see the elements that are essentials and the ones that are missing such as titles, sources, etc. Sooner or later, it is possible to present traditional graphics such as bar graphs, pie charts, line graphs and other graphical representations preparing the way to the study of functions and their graphs.

Once the data is displayed in a graph or a table, it is time to examine "what the display tells us?" (Dunkels, 1991, p. 131), these demands are important and will make all the difference, they must be interesting and rich. It may start with counting "How many…? and go on with comparisons: "Which is the largest…. Or which is the smallest…? Questions like: "Which fruit was chosen by less than 3 children? or " at least 2 or…more than 5? ask for counting but also introduce important terms "less than, at least, more than. Then it is necessary to sort to find "what would be the second choice" and so on to get to a complete order of all choices. The distinction between cardinal and ordinal numbers is also introduced.

Other questions may require addition and subtraction. In a problem where the story is about a survey done by fruit-stand owner, a question like "How many customers liked either apples or bananas?" (Mighton, 2004, p. 122) asks for a sum. If we ask: "how many more pears than bananas did the children in the class bring today?," the answer would be found by a doing a subtraction. It can be done by counting the dots with a one to one correspondence and stopping in a first step when the "bananas dots" are finished and going on by counting the remaining "pears dots" or really saying "seven less five is two" (with older children), Whatever method the children use, they are putting into action the concept of difference and not only doing a memorized computation. Questions like: "what fruit is twice as popular as peaches?" (Mighton, 2004, p. 122) may lead to a multiplication.

A simple data handling activity involves collecting, organising, categorising, recording and symbolising data. The activity activates number concepts, such as cardinal and ordinal, counting and elementary arithmetical operations. It is only a beginning.

Numbers will eventually replace the dots and words will take place of the drawings representing each category of data. Simple pictograms will prepare the way for organised tables where the categories will be replaced by numerical data and by intervals. The concept of number as a measure will appear, for height or weight for example. The data being quantitative, order becomes essential. Furthermore, contexts can be complicated and data will eventually draw on bigger numbers, real, rational, decimal numbers, asking for the related arithmetical operations.

STEM AND LEAF PLOTS

The stem and leaf plots used with elementary school children of various grades, using the data supplied by the ages of the mothers, help make the number concept more explicit. The children write a two-digit number and are compelled to distinguish the units and tens. Making a stem and leaf plot also asks for ordering. And when the data are expressed with larger numbers, the question of rounding will rise (how and how much). In a group of Grade 5 working with the length of the rivers of their country, the children rounded the numbers of kilometres to the 10^{th} while putting together a stem and leaf plot (Dunkels, 1991).

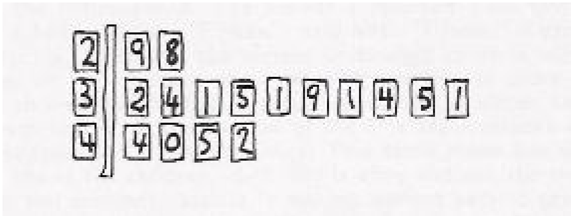

Figure 2: Dunkels (1986 p. 63)

In addition to the mathematics previously identified, the questioning here again can activate other reasoning and concepts. In this case, data are quantitative and allow new questions: “The older half of the mothers is older than what age?,” “It is possible to look for the data that separates the group in two equal groups?”. And then separating each new group in two, thus separating the original group in four parts, it permits not only to identify the median and the quartiles, but also to work on rational numbers -half and quart- and on ordinal numbers.

At this point, characteristic measures can be introduced. The computing algorithm for the mean uses addition and division in the sense of “equal sharing,” it is a “shared total” but the resulting quotient can also be seen as a ratio, for example “2,1 children/family.” The variability concept clearly calls for a difference and the notion of greater and smaller. These concept can be exploited not only to exercise addition and division but overall to give sense to arithmetical operations.

PROPORTIONALITY

Various studies (René de Cotret, 1991; Pézard, 1986) show that proportional reasoning and, particularly, percentages are not well mastered. It was found that lyceum students and even student teachers still had difficulty with percentages (Buisson, 1981; Carayol, 1983). The concept of proportionality is a basic mathematical concept that encloses that of percentage but is also fundamental for understanding the linear function. It is also linked to fractions and decimals and graphical interpretation (Pézard, 1986). One of the reason is that often proportionality is presented as a “règle de trios,” so there is not really a gain of knowledge but a “training to follow certain rules” (Pluvinage and Dupuis, 1981). Pluvinage *et al.* (id.) presented problems of proportionality in geography, physics or mathematics to 13 year old pupils and found that the results were coherent independently of the subject. Rouchier (1980) suggests that it is through problem solving that a pupil gets to master a concept and that instruction should not separate mathematics from its applications.

Here again, statistics offers advantageous opportunities to work on all these concepts. Many daily life situations, as objects and their price, speed and time, can be subject of a statistical survey and will deal with the foundation of the proportionality concept, namely the multiplicative structures. For example, in an activity proposed in the Schools Council Project on Statistical education (11-16), (1980), it is suggested to work on “Leisure time”: “1-What do you do in your spare time?” “2-How long do you spend watching television?” “3-How do your figures compare with the national figures?” (p. 93). The data associated to the first question could be represented by a pie chart. This will require the use of proportion as a fraction: 7 out of 22 students play soccer (7/22). Next, this fraction has to be transformed in a equivalent fraction of 360°, an operation that requires proportional reasoning. Preparing pie charts gives an opportunity to work with fractions seen as part of a whole, carrying out multiplication to find the portion of 360° that will represent a particular characteristic and to draw the corresponding sector of the pie. It certainly presents an occasion to work on proportional reasoning and equivalent fractions, their equivalent percentage or vice-versa.

A bar chart may also be constructed, adding the necessity to deal with scales. Here again there is a proportional reasoning, the frequency being proportional to the length of the bar. These operations are too often considered as trivial even if they are an essential step of mathematical development if seen from the reasoning point of view instead of the procedural operational one. Constructing relative frequency tables will also put into use the same types of reasoning.

If we go on with the second question, time looking at television, it is necessary to cope with a continuous variable and to introduce class intervals (although it is always possible to translate in a discrete variable for younger pupils). There arrive “open or closed intervals” and the choice of their width. This will conduct to constructing histograms that also requires proportional reasoning, here for a portion of an area, especially if the data are reported in intervals of different width. The important aspect, at this point, is the proportional link between the frequency, expressed as a number or fraction or percentage, and the area which is a two dimension measure. This is a crucial step in the construction of the concept of distribution functions.

At this point, it is possible to introduce average and dispersion measure for grouped data. Since the raw data is available, it is possible to compare the results obtained with them to the ones found using the middle value of the interval, thus leading to a discussion on estimation.

To compare their own figures with national figures (Question 3), the students need to convert the frequencies into percentages. In a experiment using “Recensement à l’école” (CensusAtSchool: http://www19.statcan.ca/) in a 7th grade class in Montreal, it was observed that to compare their data to the one found on the website for the whole country, students tackled with much more complex values than the ones usually found in exercise books for finding percentages But they were quite successful and surely highly motivated. This work may be drawn on to give a more profound insight into percentages, examining the relative modification of the percentage for different totals. For instance, for a group of 20, two more is 10% more, but if the group size is two thousands, what difference in the size will generate a variation of 10%? And so on.

Box-plots are also helpful for comparisons and will accentuate the work on the number line, the concepts of half and quarter but also of 25 %, 50%, 75% and variability. It is also another important place to work on the distinction of the role of numbers: data, frequency, relative frequency. This is easier when numbers are presented in a real context. Taking advantage of multiple pictorial forms to analyse the data is important, since graphical methods seem to stimulate thinking, contrarily to formulae that often rely on automatism (Vännman, 1990).

VARIABILITY AND FUNCTIONS

Other cases of mathematical concepts are variability and functions; these concepts are widely used in statistics. For example, line graphs representing the relationship between two variables can be used to observe trends and to make predictions, another way to interpret numbers and eventually for modeling. It is also an introduction to the study of variability and functions. This is as well true for linear relations and regression. Scatter plots can be introduced quite early (Mighton, 2005). Doing so, children get familiar not only with scales, but also with coordinates and again with relations between variables. In high school, it is possible to look into linear relations. In a first step, an estimated straight line can be drawn and from there, we can work on finding the equation from given information such as two points or the y-intercept and the slope.

In a further step, students can look for the straight line (or eventually, the shape that will be optimal to translate the relation between both variables. In doing so, they must use the concept of distance between a point and a line or between two points and handle absolute values and absolute value functions. It will, then, be necessary to deal with sums of distances or of their squares. The study of both absolute value and quadratic functions, particularly when looking for the minima of these functions, is needed to explain why the sum of squares of the differences between the y-coordinate of a point and the corresponding y-coordinate on the line was preferred to the sum of absolute values of the differences in order to find the “best” line.

OTHER EXAMPLES

And we did not mention technology. From elementary to high school, the study of statistics is an excellent opportunity to use calculators (graphic or not), computers and the Internet. Not only, the student familiarises himself with the use of these “machines but he also needs to develop a critical spirit in front of the multiple productions. In addition, technology leaves space to the development of imagination and creativity, questioning, analysing and interpreting. “Does this graphic make sense?” (Huff, 1954) or “the chi-square value is…” “What does it mean?” If pupils question if their choice of pets is similar or different from the ones of children living in other countries, they will find the data on Internet, for this and numerous others characteristics, since a wide choice of data bases are freely available for education institutions.

DISCUSSION

Mathematics is imbedded in statistics and statistics–is an ideal field to provide a meaningful context for the learning of many mathematical concepts. Moreover, doing so, students not only intensify their understanding of mathematics but they also get in touch with another discipline that occupies more and more place in their everyday environment. They get to know another point of view where answers are relative or not straightforward but arguments are based on quantitative results. “Statistics should not be taught as a separate unit, but should be introduced whenever appropriate to illustrate and expand upon standard concepts (such as measurement) and to form interdisciplinary links for students” (Burrill, 1990, p. 222). “… statistics can be involved in the mathematics course without necessarily “throwing out” important mathematics areas. Instead mathematics is strengthened by discussions around statistical matters, and using statistics we can put reality into the mathematics courses” (Vännman, 1990, p. 120).

Nevertheless, the aim of our demonstration was primarily to contest the belief that statistics are “easy” and that it takes too much of that precious time needed for the mathematic curriculum. The study of statistics can be fully integrated in the mathematic curriculum giving it meaning related to everyday life. Pupil curiosity about real life can foster interest in mathematics as a tool to understand reality (Dunkell, 1990).

Finally, statistical activities in the classroom can be directly connected to the students personal interests and stimulate their motivation for numerical and quantitative studies.

ACKNOWLEDGEMENTS

Thanks to Prof. Miriam Pannone for her helpful suggestions.

REFERENCES

— (1980). *Schools Council Project on Statistical Education* (11-16), Berks, England: W. Foulsham and CO. Ltd.

Aksu, M. (1990). Problem areas related to statistics in training teachers of mathematics in Turkey. In A. Hawkins (Ed.), *Training Teachers to Teach Statistics Proceedings of the International Statistical Institute Round Table Conference*, Budapest, Hungary, (pp. 127-137). Voorburg, The Netherlands: International Statistical Institute.

Buisson, P. (1981). *Pourcentages: Appropriation – Transfert.* IREM de Rouen.

Burrill, G. (1990). Quantitative literacy; leadership training for master teachers. In A. Hawkins (Ed.), *Training Teachers to Teach Statistics Proceedings of the International Statistical Institute Round Table Conference*, Budapest, Hungary (pp. 219-227). Voorburg, The Netherlands: International Statistical Institute.

Carayol, F. (1983). *Comportements d’élèves et de futures maîtres de l’école élémentaire face à des questions de mathématiques.* Thèse de 3e cycle. IREM de Toulouse.

Cobb, P. (1999). Individual and collective mathematical learning: The case of statistical data analysis. *Mathematical Thinking and Learning*, 1, 5–44.

Darrell, H. (1954). *How to Lie with Statistics*. New York: W. W. Norton.

Dunkels, A. (1990). Examples from the in-service classroom (age group 7-12). In A. Hawkins (Ed.), *Training Teachers to Teach Statistics Proceedings of the International Statistical Institute Round Table Conference*, Budapest, Hungary (pp. 102-109). Voorburg, The Netherlands: International Statistical Institute.

Dunkels, A. (1986). EDA in the primary classroom. Graphing and concept formation combined. In R. Davidson and J. Swift (Eds.), *Proceedings of the Second International Conference on*

Teaching Statistics, Victoria, B. C.: University of Victoria, (p. 61-66). Voorburg, The Netherlands: International Statistical Institute.

Dunkels, A. (1991). Interplay of the number concept and statistics using attitudes and techniques of EDA. In D. Vere-Jones, S. Carlyle and B. P. Dawkins (Eds.), *Proceedings of the Third International Conference on Teaching Statistics*, Dunedin, New Zealand, Vol.1, (pp. 129-139). Voorburg, The Netherlands: International Statistical Institute.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3), www.amstat.org/publications/jse/v10n3/garfield.html.

Gattuso, L. and Pannone, M. A (2001). Una sperimentazione di didattica della statistica e gli insegnanti che l'hanno condotta. *Induzioni*, 23, 185-204.

Gattuso, L. and Pannone, M.A (2002). Teacher's training in a statistic teaching experimentation. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*, Cape Town, (pp. 685-692). Voorburg, The Netherlands: International Statistical Institute.

Gattuso, L. and Pannone, M. A. (2000). *Une expérimentation d'enseignement des statistiques et les enseignants qui l'ont vécue*. Sperimentazione di nuove strategie didattiche per l'apprendimento della statistica cofinanziata dal M.U.R.S.T. e dalle Università di Padova, Palermo, Perugia e Roma "La Sapienza" (1998-2000): Giornate di Studio, ISTAT, Roma.

Hawkins, A., Jolliffe, F., and Glickman, L. (1992). *Teaching Statistical Concepts. The Effective Teachers Series.* London: Longman.

Konold, C. and Higgins, L.T. (2003). Reasoning about data. In J. Kilpatrick, W.G. Martin and D. Schiffer (Eds.), *A Research Companion to Principles and Standards for School Mathematics,* (pp. 192-215). Reston, VA: NCTM.

Konold, C., Pollastsek, A. and Gagnon, A. (1996). Students analyzing data: Research of critical barriers. In J. B. Garfield and G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics: 1996 Proceedings of the 1996 IASE Round Table Conference* (pp. 151-167). Voorburg, The Netherlands: International Statistical Institute.

McGatha, M., Cobb, P., and McClain, K. (1998). An analysis of students' statistical understandings. Paper presented at the Annual Meeting of the American Education Research Association. San Diego, CA.

Mighton, J. (2004). *JUMP at Home—Grade 3: Math worksheets for the elementary curriculum.* Toronto: House of Anansi Press Inc.

Mighton, J. (2005). *JUMP at Home—Grade 6: Math worksheets for the elementary curriculum.* Toronto: House of Anansi Press Inc.

Pézard, M. (1986). À propos de l'enseignement de la proportionnalité . Les cahiers de didactique. Num. 20. IREM de Paris7, Paris.

Pluvinage, F. and Dupuis, C. (1981). La proportionnalité et son utilisation. *Recherche en didactique des mathématiques*, 2(2).

René de Cotret, S. (1991). *Étude de l'influence des variables indice de proportionnalité du theme et nombre de couples de données sur la reconnaissance, le traitement et la compréhension de problèmes de proportionnalité chez des élèves de 13-14 ans.* Thèse présentée à l'Université Joseph Fourier, Grenoble I.

Rouchier, A. (1980). Situations et processus didactiques dans l'étude des nombres rationnels positives. *Recherche en didactique des mathématiques*, 2(2).

Russell, S. (1990). Counting noses and scary things: Children construct their ideas about data. In D. Vere-Jones, S. Carlyle and B. P. Dawkins (Eds.), *Proceedings of the Third International Conference on Teaching Statistics*, Dunedin, New Zealand, (pp. 158-164). Voorburg: International Statistical Institute.

Russell, S. (1990a). Issues in training teachers to teach statistics in the elementary school: A world of uncertainty In A. Hawkins (Ed.), *Training Teachers to Teach Statistics Proceedings of the International Statistical Institute Round Table Conference*, Budapest, Hungary, (pp. 59-71). Voorburg, Netherlands: International Statistical Institute.

Vännman, K. (1990). Some aspects of statistical graphics for secondary school teachers. In A. Hawkins (Ed.), *Training Teachers to Teach Statistics Proceedings of the International Statistical Institute Round Table Conference*, Budapest, Hungary, (pp. 110-125). Voorburg, Netherlands: International Statistical Institute.