STATISTICS VISUALIZATION WITH DYNAMIC GEOMETRY

José Alexandre dos Santos Vaz Martins
Instituto Politécnico da Guarda, Portugal
jasvm@ipg.pt

*Students from this generation, learn best by seeing. Nowadays, visualisation is extremely important to help young students to catch the real meaning of some concepts. Therefore, this paper explores the application of a dynamic geometry software (*Cabri-Géomètre II*) to illustrate basic statistical concepts (like mean, median and mode), their properties and graphical representation, in a way that allows the teacher to explore, discover and uncover those ideas that will get the messages to students and stimulate interactive work in the classroom. In this presentation a main goal is showing a way to approach statistics through the use of a computational tool, with emphasis in the visual exploration of statistical concepts and focus on the improvement of statistical literacy, reasoning and thinking.*

INTRODUCTION

As Bento de Jesus Caraça wrote (2000), the Greeks, in the V century BC, had imposed a separation between the number and the figure. This separation between Geometry and Arithmetic, and the exclusion of Geometry from all which had some connection with a quantitative and dynamic conception, came in the XV and XVI centuries. But, with the end of this separation Analytic Geometry was born, which finished the superposition of the number over the figure. So, the proximity of geometric and analytic fields, in the last four centuries, has been a significant issue in the history of knowledge.

Despite that and according to Miguel de Guzmán (2001) even during a great part of the XX century strong formal tendencies existed that sent the relation between the number and the figure to a secondary role, and consequently a certain apprehensive and suspicious regard to mathematic visualisation.

In spite of that, a recent trend from that clearly emerge visualisation with a brand new role in mathematic and statistic teaching. This legacy allows us, today, with the computers help, to go further, as learners as well as teachers.

The truth is that, in Mathematics and in Statistics, some ideas, concepts and methods present a great deal of richness of visual contents.

So, like Miguel de Guzmán (2001) refers about this topic, it is natural to consider visualisation as an extraordinary important aspect in mathematic general activity as task of creation, discovery of new relations or of knowledge transmission.

From this point of view, *Dynamic Geometry Software* (*DGS*) could be very useful, because they have constructive, measurement and dynamic proprieties that allow the creation of some applications with a huge potential in terms of the visualisation characteristics that were mentioned.

Sometimes, with *DGS* assistance, it is possible to help students overcome some of their difficulties, for instance: grasp of knowledge that implies logical and intuitive thinking; systematisation; argumentation; interpretation; selection and evaluation of information value; development of a multidisciplinary attitude.

These statements are a truth that we can apply to numerous mathematic fields like complex numbers, differential equations, function analysis, numerical analysis, trigonometry, and also to statistics.

So, the intention of this paper is to present the *DGS* potential at the service of the Descriptive Statistic basic concepts apprehension, contributing, in this way, to improve statistics teaching, that has reached a overwhelming importance in a curricular and social perspective. It is also intended to facilitate the teacher's work showing ideas, trails and new ways of reaching such objectives. As Godino (1995) suggest, these software programs don't solve teaching problems by themselves, but it is necessary a great deal of reflection and investigation to make adequate didactical guides to teach specific concepts.

In this sense, some *Cabri* applications built to get a visual stimulation and to facilitate mean, median and mode concepts acquisition, trough a geometric interpretation of them and of theirs proprieties will be presented, hence dynamic manipulation and exploration of geometric properties may play an important role in teaching and learning statistical concepts and to experience the effect of data values changes on statistical measures. Some *Cabri* applications with some type of graphics to explore and to establish interrelationships will also be present.

APLICATIONS PRESENTATION

All of the applications presented in this paper are constructed in the dynamic geometry software, *Cabri Géomètre II*, and most of them could be implemented with rudimentary knowledge of *Cabri* tools.

After this introduction the referenced issues and some attached activities will be presented.

Median

According to Cobo and Batanero (2000) the usual median definitions present a few difficulties and ambiguities that persist in the calculus work. The same authors state that the median calculus is based in cumulative frequency display comprehension and in a correct proportional reasoning by the students and also in the understanding of triangle similarity concept.

For the median, the application purposed shows how to get it from an cumulative frequency display (the *x* corresponding to the 50% *y*), establishes the relation between the histogram and the cumulative frequency display and, based in triangles relations, deduces the formula for the median calculus in the continuity case, as it is possible to see in Figure 1.
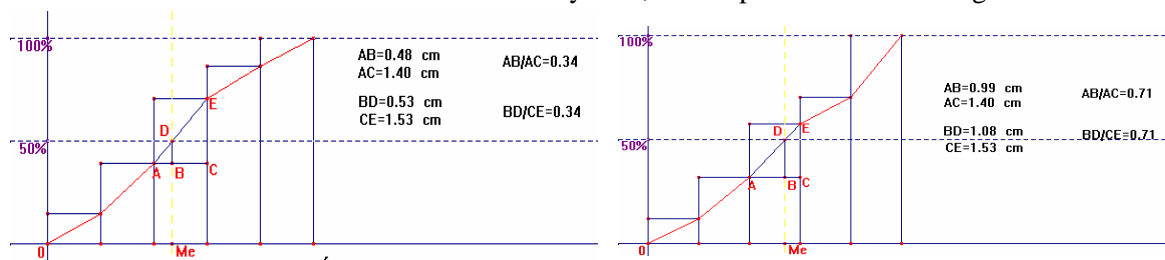

Figure 1: Geometrically obtaining the median and its formula

Besides, it is possible to explore some numerical, algebraic, statistical and geometrical properties. One example is a dynamic verification of the equality of the median right and left areas in the histogram as it is shown in Figure 2.
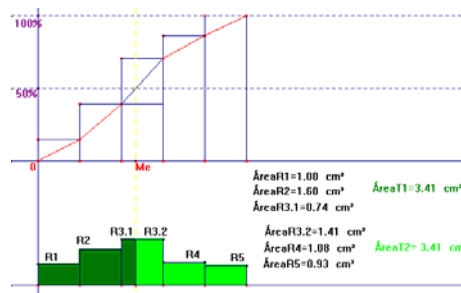

Figure 2: Equality of the median right and left area

Mode

As it is known, there are several formulas to establish a value for the mode in the modal interval. This value is an approximation, and therefore needs a careful interpretation. In this sense, it is very important to have a better knowledge about how those formulas work.

Based in the principle that, in the modal interval, the mode should be closer to the contiguous interval with the greatest frequency now, like it is visible in Figure 3, two geometric processes will be approached to obtain the mode value and its formula of calculus.
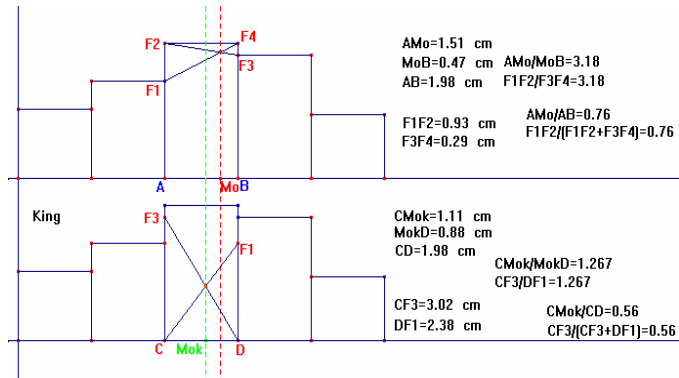

Figure 3: Geometric processes to obtain the mode

From a duplicated histogram in a *Cabri* application, it is possible to compare the values of the mode obtained by the two different ways mentioned. One can also study and understand the behaviour of each geometric process; especially with extreme values in the neighbours' intervals and to find in what situation both values are equal. From triangle relations, it is also possible to deduce the formulas to calculate a value for the mode and confirm that in a dynamic way.

Mean

As it is well known, the mean could have a physical interpretation as the gravity centre. So, based in the arithmetic mean property $\sum (x_i - \bar{x}) = 0$, in the *Cabri* application proposed a scale was constructed where the respective arithmetic mean represents the equilibrium point, as it is possible to confirm in Figure 4.

In this application, beside the physical interpretation, it is also possible to confirm analytically the value of arithmetic mean and it could help as a dynamic approach to some of the difficulties related to the arithmetic mean that were mentioned by Batanero (2000).
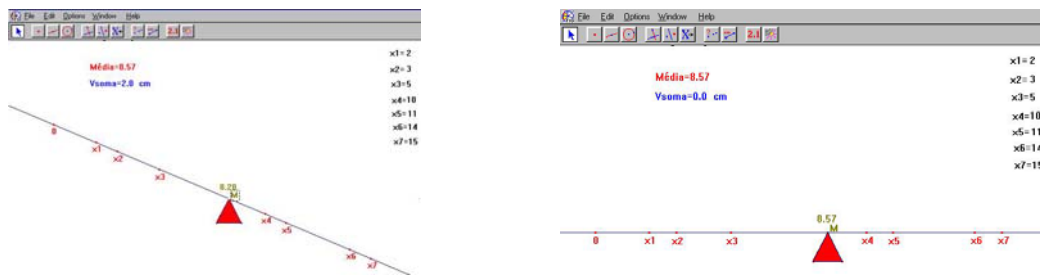

Figure 4: The arithmetic mean scale

Others applications could be implemented for values with non-unitary frequency or for a continuous variable.

Graphics

One example of the use of *Cabri* applied to graphic issues is one that remarks the relation between a bar graphic and the correspondent pie graphic and presents the constant relation between the bar areas and the respective angles as it is shown in Figure 5.
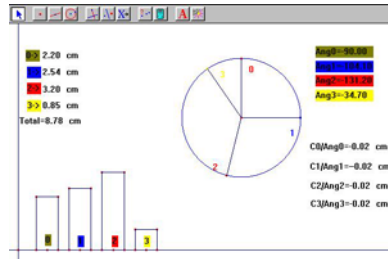
Figure 5: Relation between bar and pie graphics

Finally, other *Cabri* application dynamically shows the relation between histogram, cumulative frequency display and box plot, like it is visible in Figure 6, making the understanding and interpretation of this diagram easier, which is many times so difficult for the students.
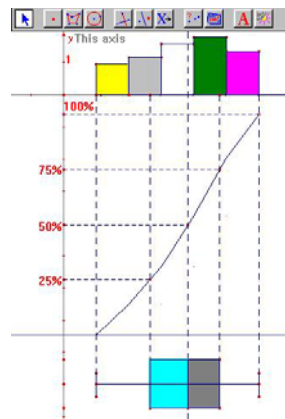

Figure 6: Histogram, cumulative frequency display and box plot

CONCLUSION

An improvement in statistics teaching using *DGS* in a conscientious, reflected and careful way is achievable.

There are many possibilities to explore, improve and add other potentialities to the presented *Cabri* applications, having curiosity, imagination and good will as motivation.

It is also significant to remark that it is possible the inter-disciplinary stimulation from the use of the presented and other *Cabri* (*DGS*) applications.

REFERENCES

Batanero, C. (2000). Dificultades de los estudiantes en los conceptos estadísticos elementales: el caso de las medidas de posición central. *Ensino e Aprendizagem da Estadística.* Sociedade Portuguesa de Estatística.

Caraça, B. (2000). *Conceitos Fundamentais da Matemática*. Gradiva.

Cobo, B. and Batanero, C. (2000) ¿La mediana en la educación secundaria obligatoria:?Un concepto sencillo? *UNO*, 23, 85-96.

Guzmán, M. (2001). *El Rincón de la Pizarra*. Pirámide.

Godino, J. (1995). ¿Qué aportan los ordenadores a la enseñanza y aprendizaje de la estadística? *UNO*, 5, 45-56.

Martins, J. A. (2003). Estadística con Geometría Dinámica. *Actas del VI Congreso Galego de Estatística e Investigación de Operacións - Aplicacións da Estadística*. Universidad de Vigo