

GeoGebra in Teaching and Learning Introductory Statistics

Theodosia Prodromou
theodosia.prodromou@une.edu.au
University of New England
Australia

Abstract: *This article discusses the special opportunities for teaching statistics that technology offers teachers who aim to provide rich learning experiences for their students. These opportunities involve automation of many activities such as quickly organising data, computing measures, and generating graphs. By automating the tasks of computing statistics or generating data, technology facilitates students' ability to visualise abstract concepts, affording an opportunity to focus on conceptual understandings and data analysis. This article also examines how GeoGebra can be integrated into the curriculum and learning process of introductory statistics to engage college students in cycles of investigation including (a) managing data (b) developing students' knowledge for understanding specific statistical concepts, (c) performing data analysis and inference, and (d) exploring probability models. Recommendations are included for ways mathematics educators can engage college learners in developing their knowledge for exploring data concepts and doing statistics with GeoGebra.*

1. Introduction

We are all familiar with the great importance of computer software availability and associated popularity in determining what statistical analyses are carried out and how they are presented. In the world of statistics itself, the spreadsheet Excel has long been in common use for creating graphical representations of data. In the world of statistics, there are several sophisticated software packages used by statisticians—SAS, SPSS, STATA, GENSTAT, STATISTICA, S and S-PLUS, and more recently R. Additionally, technologies such as Excel, TinkerPlots, Fathom, Autograph for doing and teaching statistics have become commonplace in many tertiary and some secondary contexts.

This article discusses (1) the many affordances that technological advances offer to teachers of statistics who aim to provide rich learning experiences for their students, and (2) the potential of GeoGebra in the teaching and learning of statistics and how GeoGebra, which is designed for engaging students in managing data, performing data analysis and inference, and exploring probability models, can be integrated into the curriculum and learning process of introductory statistics.

GeoGebra, interactive mathematical software used to teach geometry, calculus, algebra, and statistics, has become one of the major technological tools used by educators all over the world. GeoGebra, which can be an effective tool to teach geometry, calculus, algebra, and statistics, is of particular interest to educators with a limited budget because it is an open-source software package, available free on the Internet.

2. What does technology offer teachers of statistics?

Pratt et al. [1] consider five significant “affordances for teaching that that might accrue should these categories of software be more widely adopted” (see [1], p. 98) as follows: (1) Using representations as dynamic tools for analysis, (2) expressing personal models, (3) exploring models, (4) storing and processing real data, and (5) sharing and communicating.

Using representations as dynamic tools for analysis is the first affordance. Traditionally, graphs are used to report data, often through graphical displays and representations. Statistical software can provide users with easy-to-create graphical representations of data, encouraging deep interaction with data. Computers, hence enable representations to be used as analytical tools during an investigation instead of treating them merely as presentational tools at the end of the

investigation. Pratt et al. [1] argue that, “when graphs are used to try to make sense of data during analysis, the representations need to appeal to an intuitive sense of position, spread and outlying values” (see [1], p. 98). One example is the software called “Mini-Tools” developed by Cobb, Gravemeijer, Bowers, and McClain [2] designed to appeal to students’ intuitive notions of data distribution and numerical data to help students develop ways of thinking about distributions [3]. Multivariate data representations represented in complex visualizations might be better supported through novel digital tools. Future pedagogical theories should try to integrate these novel tools into student activities and school curriculum [4].

Expressing personal models is the second significant affordance for teaching statistics [1]. Statistical analysis involves the creation and use of models that describe the data generated from a real-world situation. Teachers of statistics might therefore try to engage their students in an activity of expressing personal models (i.e., write down/communicate to the computer or others) so that by trying to express different models they will come to see which models do a better job of capturing/explaining the inherent structure in a situation. Hence, students might embrace the full cycle of statistical enquiry that requires reality description, creation of pseudo-concrete model, mathematisation-formalisation and validation [5]. Computational tools provide students the ways to express their ideas needed to develop models that best fit the data. In exploratory data analysis, students express their own informal models for the data by looking for patterns in the data. This process is often termed expressive modelling [6].

Teachers could also use statistical software as an authoring tool in which they build models for students to explore by seeing sample data, and students can use such software as an expressive tool in which they build their own models of phenomena. According to Pratt et al. [1], exploring models is the third important affordance for teaching statistics. The process of engaging in expressive modelling [4] might provide opportunities for students to appreciate the use of key statistical concepts in particular contexts. When students are engaged in activities of expressive modelling, there is no guarantee that they are also engaged in constructing any key statistical concepts. Models can be built into the computer software and used to generate simulations that can be explored by the students, who might not perceive the model through data but perceive a real-world phenomenon through the mathematical model. Simulations have been used to help students draw connections between ambiguous statistical concepts such as epistemological perspectives for probability [6], unexplained and explained variation [7], or deterministic and stochastic causality [8]. Such approaches that promote experimentation with data and exploration of statistical distributions are termed as exploratory modelling [9].

According to [1] the fourth important affordance for teaching statistics is storing and processing real data. Digital technology enables the use of large data sets through its capacity for storage and easy retrieval of large data sets. The use of large data sets enables the analysis of data drawn from real situations that are meaningful to students. Real data sets present issues—such as missing values and other errors in the data—that are usually not present in sanitised data sets. Such issues need to be confronted since they raise important concerns about the limitations and reliability of statistical inferences that might be made.

Sharing and communicating knowledge is the fifth fundamental affordance for teaching statistics with technology [1]. Several researchers ([10], [11]) have argued that new technologies promote collaborative learning and exchange of information about statistics. For example, [11] has researched several types of wiki activities for the creation of resources that could be beneficial to teacher trainees of statistics and their future students. Amongst these activities, are collaborative writing, glossaries, discussion and reviews, statistical projects, self-reflective journals, and assessment. [11] investigated a very large data set consisting 9000 email messages that contained 30 variables for each message. This large data set was used to introduce teacher trainees to a dynamic document environment that involved a live worksheet environment that allowed students to interact with the data, adjusting inputs, altering computations, and exploring a variety of different

statistical analyses. A live worksheet with data provides users with opportunities to reflect on statistical investigations performed by statisticians. By allowing the electronic document to be an interactive environment, teachers can more effectively teach statistics and aid students in drawing better connections between problem-solving activities in statistics and real life situations.

The five affordances described above, which involve using representations as dynamic tools for analysing and expressing models, including personal models, for storing and processing real data, and for sharing and communicating, will be used to guide our analysis of GeoGebra software as a tool for teaching introductory statistics. We might expect that technology affordances as described above have changed the way teachers teach statistics and students learn statistics [1]. Technology in general and GeoGebra in particular, can help with amplifying [12] users' abilities to solve problems or reorganise the way people think about problems in Statistics and search for their solutions [13].

We will reflect on using GeoGebra [14] in teaching and learning Introductory Statistics at tertiary education.

3. GeoGebra

GeoGebra is a dynamic and interactive geometry, algebra, and calculus application, intended for teachers and students. Most parts of GeoGebra are free software, making GeoGebra accessible to those who cannot afford the proprietary packages like SAS and SPSS. GeoGebra is currently used by more than 20 million students and teachers in 190 countries. 155 GeoGebra Institutes (IGI) have been established in 85 countries around the world and it has been translated to 62 languages. GeoGebra is available on computers and mobile devices [15]. GeoGebra is not only a software package, but also a community that shares ideas and free educational materials supporting teachers and students in using GeoGebra, and working on projects related to GeoGebra.

GeoGebra was initially designed for teaching and learning geometry and algebra. Hence, it involves the following main features: dynamic geometry environment, built-in CAS, scripting, and a built-in spreadsheet with extended S. Geometrical constructions can be made using coordinates to determine the position of points, vectors, segments, lines, polygons, conic sections, inequalities, implicit polynomials, and functions. Elements can be entered and changed dynamically on screen, or through the Input Bar. GeoGebra linked the learning of geometry, algebra, and calculus by using variables for numbers, vectors, and points to view the algebraic formula, find derivatives and integrals of functions and it has a full complement of commands like Root or Extremum. Teachers and students can use GeoGebra to make conjectures and to understand how to prove geometric theorems because they conceptualise the mathematical formulae in relation to their geometrical constructions.

GeoGebra has very recently offered some tools that can be used for teaching statistics. The core of GeoGebra's use in statistics education has been based exclusively on supporting the achievement of student learning outcomes [13]. GeoGebra cannot simply be used as a statistical package for automating statistical calculations and producing graphical displays; GeoGebra can be used in creating Applets or using available applets created with GeoGebra when demonstrating statistical concepts in introductory statistics courses.

The core of GeoGebra's use in statistics education has been based exclusively on supporting the achievement of student learning outcomes [13]. The necessary technological skills that students need to develop to use GeoGebra alongside statistical knowledge seemed to be easier to use compared to the technological skills that students need to develop in order to operate statistical packages such as R and SAS since R and SAS require learning a whole scripting language. R use only commands-line interface and the user needs to create functions using scripting languages. Because R is designed for use by statisticians and data miners for developing statistical software and data analysis, the user interfaces of all the software programming languages are often difficult to learn. By contrast, GeoGebra's interface is designed for students, and includes features for ease

of use. For example, data can be copied and pasted from Excel and other Spreadsheet formats to the GeoGebra Spreadsheet View. Other examples of features that help students use GeoGebra are considered in the next section.

We will reflect on how GeoGebra could be integrated into the curriculum and learning process of introductory statistics.

4. Teaching introductory statistics with GeoGebra

GeoGebra can be used in two distinct ways when teaching introductory statistics courses:

- (1) Applets created with GeoGebra are integrated into lectures to demonstrate specific concepts. (See <http://www.geogebraTube.org/?lang=en>)
- (2) Students use GeoGebra as a software tool to perform data analysis and inference and to explore probability models.

These applets are used during lectures and tutorials to demonstrate specific fundamental concepts that are commonly difficult to understand in relation to a bivariate distribution. For example, the least squares line demonstration introduces students to the standard approach to the approximate solution of sets of equations in which there are more equations than unknowns. The applet of least Squares line demonstration enables students to interact with the main characteristics of the “least squares” by dragging the scatterpoints in a graph window or the two points on the fit line, or overlaying the graph with statistical measures (e.g., residuals, squared residuals, means, a least squares line) and observing the augmentations of the graph. Users can change the data by dragging points in the graph, and the relevant statistical measures update automatically.

Augmentations can be made to an existing representation to (a) alter representations, and (b) overlay characteristics other than statistical measures and additional information in the representation (e.g., a residual plot to show a graph of residuals), helping to visualise how the overall minimizes the sum of the squares of the errors made in the results of every single equation (Figure 4.1). This visualisation (Figure 4.2, Figure 4.3) is not possible without technology that can scaffold students to reorganise their conceptions of bivariate distributions.

Another applet demonstrates how the Normal distribution can be used to approximate the Binomial distribution¹. In general, if the random variable X follows the binomial distribution with parameters n and p , we write $X \sim B(n, p)$. The Binomial distribution has the following probability mass

$$f(k; n, p) = \Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

function, where

is the probability of getting exactly k successes in n trials, for $k = 0, 1, 2, \dots, n$, where

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$

This applet provides students the ability to control n and p (by moving the sliders of n and p) and to visualize the binomial distribution by observing how the altered graphical representations of the normal distribution and the binomial distribution approximate each other. The graph produced by the applet shows the normal approximation curve (and how this approximation breaks down for large or small p). If n is large enough, then the skew of the distribution is not too great. In this case a reasonable approximation to $B(n, p)$ is given by the normal distribution. If n is large enough, then the skew of the distribution is not too great. In this case a reasonable approximation to $B(n, p)$ is given by the normal distribution.

$$N(np, np(1 - p))$$

¹ This applet was made and shared by Micky Bullock <http://www.mickybullock.com>

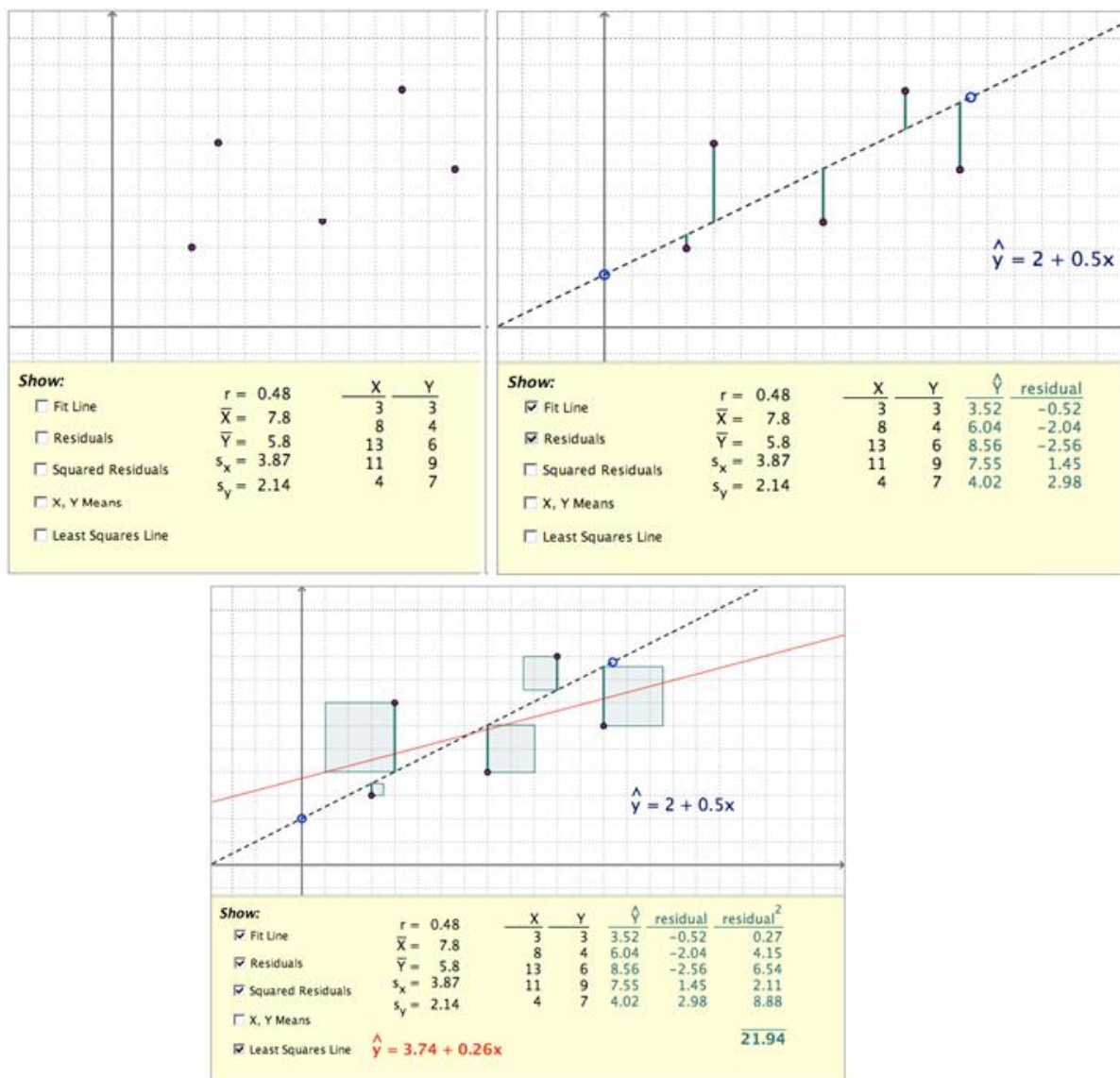


Figure 4.1 The applet of least squared line demonstration

The basic approximation generally improves as n increases (at least 20) and is better when p is not near 0 or 1. The user can also see the critical regions at either end by turning the bars red instead of green (Figure 4.4 above) – this feature is purely for visualising critical regions when performing hypothesis tests with the binomial distribution. The critical regions, however, have to be calculated by the user. In Figure 4.4 (below), the normal approximation is shown as a black line for comparison with the binomial distribution. Figure 4.4 also illustrates how to create dynamic links between the binomial distribution and the normal distribution for different values of p . In fact, p is animated to take different values that in turn animate the graphical representation of the critical regions of lower and upper end of the binomial distribution.

Linear Regression - Least Squares

An applet for understanding linear regression - an important statistical concept in data analysis.

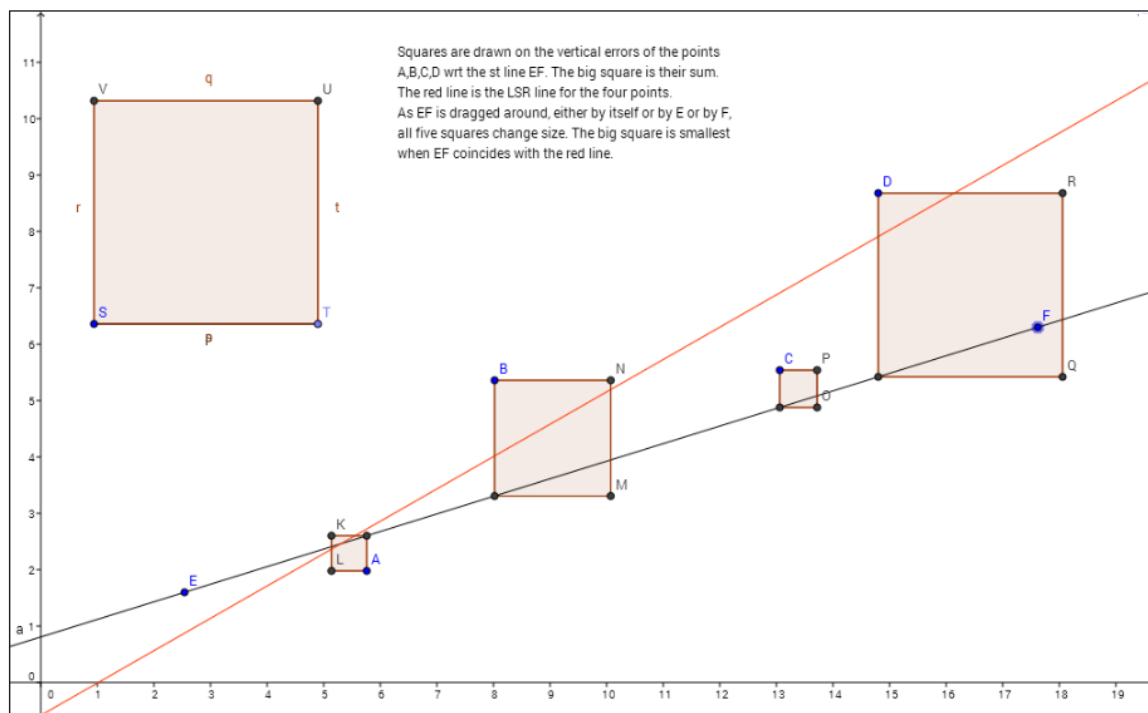


Figure 4.2 The applet of the liner regression-least squared line demonstration
(See <http://www.geogebraTube.org/student/m14559>)

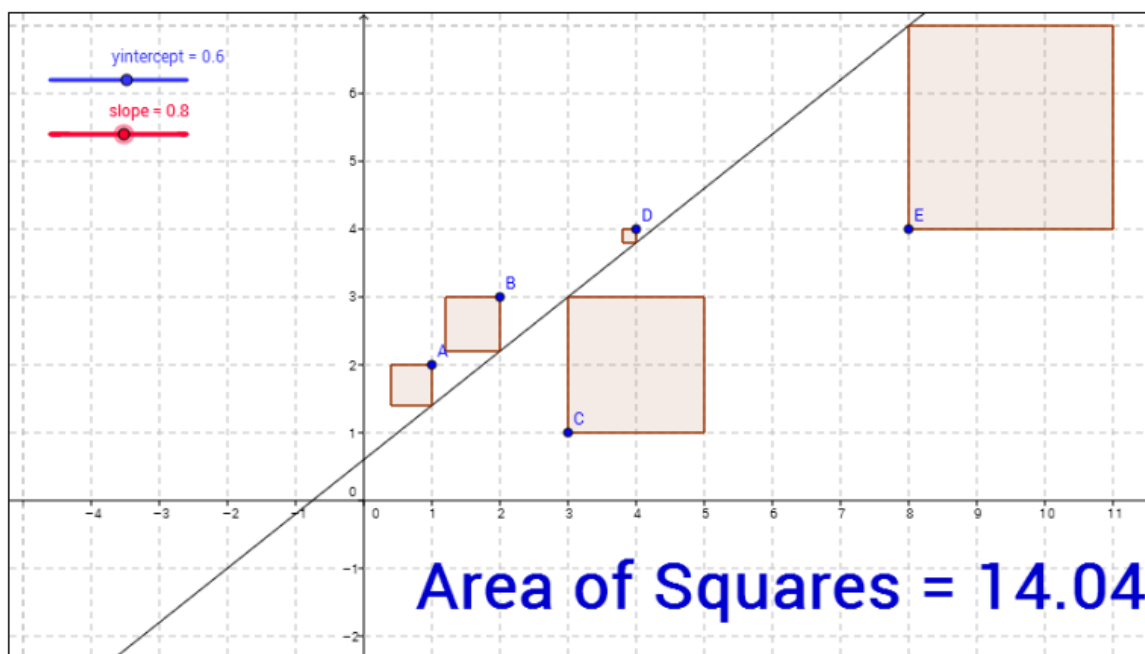


Figure 4.3 The applet of the least square regression (See <http://www.geogebraTube.org/student/m27447>)



Figure 4.4 The critical regions at either end are indicated by the red bars (above). The normal approximation is the black line overlaid on the binomial distribution (below).

A dynamic link, in contrast to a static link, occurs when a purposeful action is done to one representation that causes a reactive and coordinated action in another representation [16]. A static link occurs when there is an indication of coordinating complementary information in two or more representations [17] that does not involve any direct technological interaction with a representation.

Ultimately, GeoGebra applets like those shown above provide affordances of dynamic statistical software such abilities as (1) relating the two representations, (2) connecting them dynamically, and (3) obtaining simultaneously graphical augmentations of the animated p values, the critical regions and the normal approximation of the binomial distribution. By taking advantage of these affordances, users could actively be engaged in statistical problem solving activities that may lead to further investigations or additional insights.

In this study we are also interested in how teachers can take advantage of the features in GeoGebra that allow some of the affordances described by [1] in their teaching of introductory statistics to engage college students in (a) managing data, (b) performing data analysis and inference, and (c) exploring probability models. A description of (a) to (c) follows.

5. Using GeoGebra as a statistics software package

GeoGebra can be used as a statistics software tool to help students perform data analysis and inference and explore probability models. Students can work with a specially configured version of GeoGebra designed for a statistics-learning environment (Figure 5.1). This specially configured version of GeoGebra is only accessible by the students of the lecturer² who has designed this version of GeoGebra.

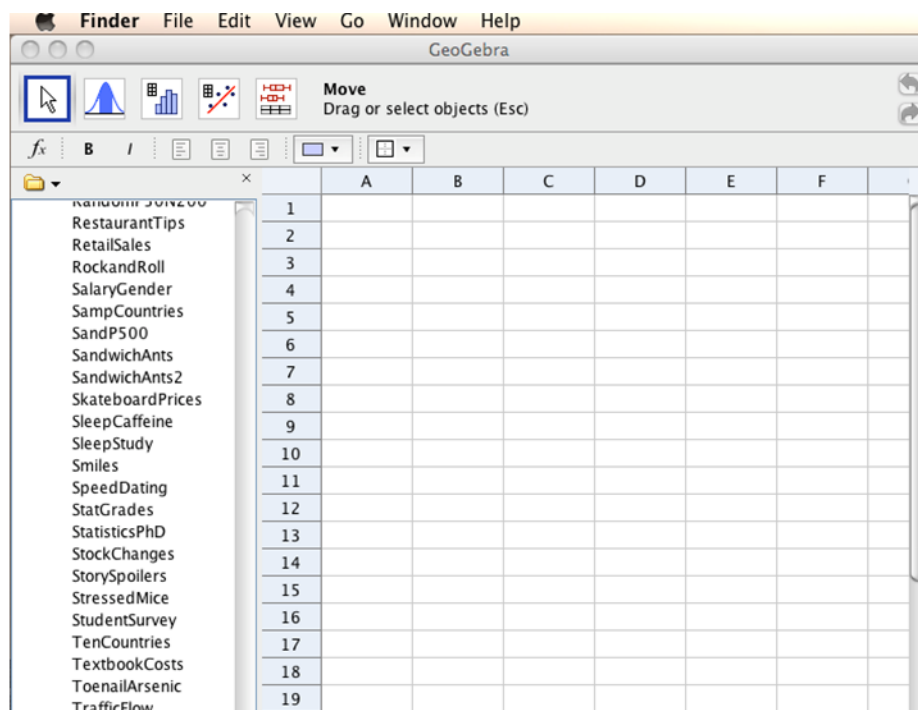


Figure 5.1 The version of GeoGebra

In this example, the students do not use the input bar or graphics window. All work is done from the spreadsheet and data analysis dialogs. This action of using sets of data points from a spreadsheet and data analysis dialogs is the primary building tool of GeoGebra when creating the representation of data analysis tools to establish a link between the data analysis graphs or summary tables to the data that is stored in the spreadsheet table. The users of GeoGebra view a spreadsheet table of data points and a graphical representation separately. This capability of storing and processing real data reflects the fourth affordance described by [1]. Other affordances described by [1] are exhibited, including the use of graphical representations for analysing data (first affordance) while students express models that describe the data (third affordance), including expressing personal models (second affordance).

A student may link the information from the data points in the spreadsheet table with the information in each plot. This way of coordinating information is done by establishing a static link between the data points of one variable and the information in each plot or summary table. The structure of the GeoGebra environment does not allow representations of data to be linked as they are created as other dynamic environments do (e.g., Fathom or TinkerPlots). In Fathom [18] or TinkerPlots[2], when creating a graphical representation or a summary table, one can drag an attribute name (variable) from the data card or data table onto a plot window or an axis in a graph window or a row or column in a summary table.

² George Sturr has designed this version of GeoGebra.

5.1. Managing Data

Data sets are re-entered and manipulated as columns in GeoGebra's spreadsheet. Data can be imported into the spreadsheet with cut and paste techniques or by opening data files from a local drive or external web site. Accessing data files is greatly simplified with GeoGebra's data browser, which displays a catalog of locally stored data files or a catalogue of files stored on a web page. With one click on a file title the spreadsheet is cleared and the data file imported. In Figure 5.1, the data browser is set by default to list the data files for the course textbook examples and homework problems.

An optional sidebar feature allows students to quickly open and explore data files. Storing and processing data as well as analysing and displaying data have the features of the four affordances as described by [1]. GeoGebra also facilitates the use of large data sets of interest to students that are collected and accessed through the Internet. Knowing how to collect real data from the Internet and how to transform it into usable data in a particular piece of software (e.g., GeoGebra) is a useful skill. The use of discussion forums, wikis, interactive whiteboards, and self-assessment software can promote collaboration and students' involvement, supporting a focus on collaborative tools for data collection, analysis, and data visualisation. Collaborative tools to collect data can promote a community approach to generating and analysing data and foster sharing of information, collaborative investigation of data, and discussion of the statistical problems. This reinforcement of sharing and communicating constitutes the fifth affordance described by [1].

5.2. Creating Graphs and Performing Inference

GeoGebra provides students with opportunities to create graphs as tools for analysis or inference. To make inferences using GeoGebra, a student selects data from spreadsheet columns and presses a toolbar button to launch a data analysis dialog window. The dialog opens with a simple graphical representation of the data and offers options to display statistical measures and sub-dialogs for calculating confidence intervals and hypothesis tests (Figure 5.2).

There are three data analysis options listed as follows: One Variable Analysis, Regression Analysis, and Multi-Variable Analysis.

The One Variable Analysis button is used for single variable data, both quantitative and categorical. It generates bar charts, histograms, boxplots, stem and leaf plots, dot plots, or normal quantile plots. The Regression Analysis button is used for quantitative bivariate data and generates scatterplots or residual plots with a variety of regression models. The Multi-Variable Analysis button generates stacked boxplots and one-way ANOVA tables.

For some dialogs there is a data-selection panel for setting data values. This option often affords opportunities to consider a view of a distribution of a group of data values and to juxtapose an aggregate view of a distribution. For example, the students might use graphical representations to consider the impact of particular data values in a distribution of aggregate values or to examine the impact of removing an outlier from a data set on measures of center. GeoGebra also allows students to focus on an interesting subset of data by allowing them to segment a distribution's range into several parts, and group data visually within a range in order to examine special features in subregions of a distribution.

5.3. Probability and Statistical Inference Calculator

GeoGebra provides students with a Probability Calculator (Figures 5.2 and 5.3). Students can use this feature to explore data, calculate probability models, and make claims and decisions about basic statistical inferences from data. The software tools in this dialog are equivalent to or surpass those found in a graphing calculator.

The Probability Calculator within GeoGebra has two tabbed panels, Probability and Statistics. The Probability tab panel allows students to view a variety of common probability

models. In particular, it displays a graph for a selected model and has a panel with a simple, button-driven interface for computing interval probabilities (Figure 5.2). The graph gives visual feedback

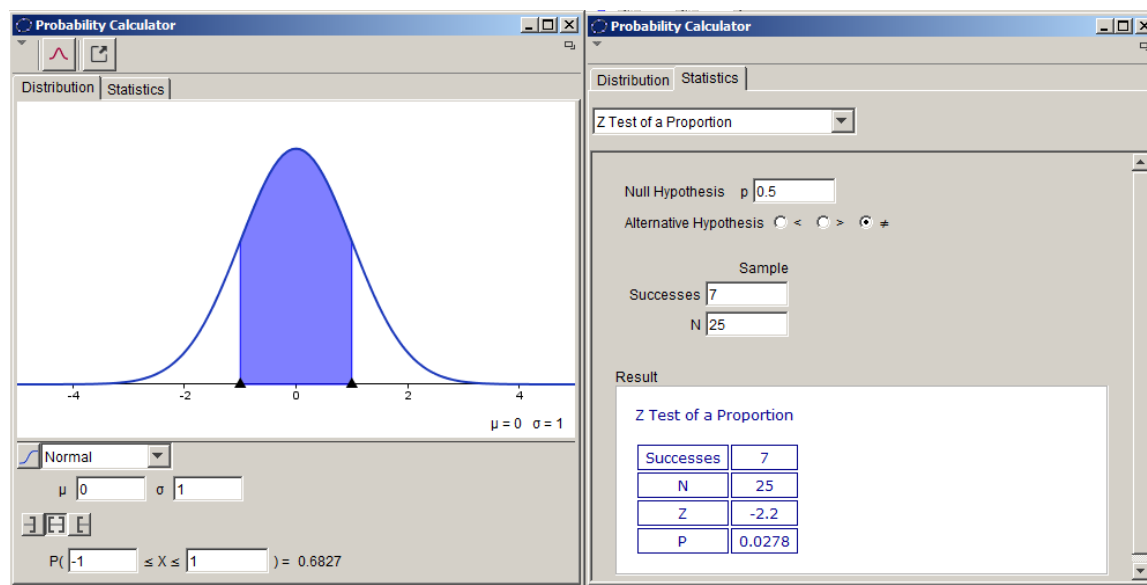


Figure 5.2 Distribution Tabbed panel (left) and Statistics Tabbed panel of Probability Calculator (right)

for computed probability intervals and allows for dynamic changes by dragging interval endpoints. This dynamic dragging of interval endpoints allows students to establish a dynamic link, coordinating the graphical representation of a distribution with statistics like mean, variance, and standard error, as well as tools for different hypothesis tests. Students are able to use dynamic linking to look at basic probability models and make inferences from data by performing hypothesis testing and observing every time how the values of mean, variance, and standard error are modified (Figure 5.2). Students can compare the impact of modifying a certain model to a statistic, and using such comparisons, they can make statements about relationships.

Students establish dynamic links when coordinating information about graphs of a selected model and computed probability intervals, which allows for dynamic changes by dragging interval endpoints. Students can make static links near the end of the process of making claims and decisions about basic statistical inferences from data, and they can reflect on their interactions with the tools of the probability and statistical inference calculator.

The Statistics tab panel is a dialog for computing common confidence intervals and hypothesis tests from summary statistics. These software calculator tools replace table lookups and hand-held calculators.

6. Discussion

GeoGebra provides the five affordances that Pratt et al. [1] suggest that technological advances offer to users. In GeoGebra, representations of data are used as dynamic tools for analysis at the end of each investigation when solving statistical tasks. This feature of GeoGebra matches the first affordance described by Pratt et al. [1]. The representations that are available in GeoGebra however, are lacking three important structures of dynamic environments. First, the graphical representations are created without establishing a connection between the summary table and the graph to the data that is stored in the data table. Hence, the graphical representations of data (e.g., tables, graphs) are not associated with each other while they are constructed. When one creates a graphical representation or a summary table, he/she cannot drag a name of a variable from the data card or data table to drop it onto an axis in a graph window or a row or column in a summary table.

Second, GeoGebra does not yet provide tools that could have an appeal to students' intuitive sense of some statistical concepts. It would be better for middle school students and younger users of the software to have access to individual data values so they could interact directly with them. For instance, dragging a data point within a dot plot could contribute to instantly separating data values into groups for the same variables and generating common graphical displays of those variables. Students could also actively construct different representations by dragging the name of the representation on to a selected representation. Third, in GeoGebra when highlighting a data point in one representation, it does not create a highlighted display of that data point within all other representations. GeoGebra does not yet include built-in connections among all representations created from a data set.

GeoGebra affords users with the fourth affordance pointed out by [1], in that it facilitates real data storage and processing of real data drawn from situations that are meaningful for the students. It can be used by students to process data quickly, calculate sophisticated statistics, analyse data, and draw appropriate inferences. The way that GeoGebra supports users to explore probability models places emphasis on the third affordance of modelling as well as on data analysis as discussed above. This feature of GeoGebra also matches Pratt's fifth affordance of sharing and communicating statistical results through discussion forums, wikis, interactive whiteboards, and self-assessment software to promote collaboration and students' involvement.

The Probability tab panel in the probability calculator could help users to view several graphical representations of different probability models, offering students a range of affordances that could scaffold their learning of statistics, ensuring that students could link statistical thinking to probabilistic thinking.

GeoGebra is a good example of software that has the affordances described by [1] and could play a key role in facilitating the development of students' understanding of statistics. The development of statistical tools in GeoGebra, will of course continue, and no doubt so will the careful analysis of how its features can be used to facilitate mathematical teaching and learning. Such developments of mathematical software need to be examined and evaluated to help mathematicians and software designers better understand the consequences of learners' interaction with the software, and ways that teachers can integrate the special characteristics of GeoGebra into teaching to support students' learning of statistical ideas. Nevertheless, the studies that focus on understanding the development of statistics learning when using GeoGebra, are limited (See [19]).

In general, observing a group of 50 University students (15 students study Bachelor in psychology, 15 students study Bachelor in mathematics and 20 pre-service secondary mathematics teachers) interacting with GeoGebra, SPSS, SAS and R, and discussing with them the impact of these statistical packages or software on their learning of introductory Statistics, provided thought-provoking insights into how GeoGebra and other statistical packages are perceived by students.

The volunteer students, who study Bachelor in psychology, attended training sessions in GeoGebra and SPSS. Almost all students who study Bachelor in psychology appeared to have acknowledged the importance of learning SPSS and recognised that learning to use SPSS has greater-long term value since SPSS is commonly used to analyse data analysis projects in social sciences.

Most of these participants commented that it would have been more time effective to learn to operate the necessary skills and dispositions required to master SPSS package instead of learning to operate GeoGebra software because GeoGebra is more suitable for introducing school students or first Year pre-service teachers to statistics. Nonetheless, these students appeared to have not appreciated the immediate relevance of SPSS package skills required to operate SPSS. They asked whether SPSS training skills could be relevant to their current studies and they referred to the future applicability of SPSS in their careers.

Almost all participants, who study Bachelor in psychology, expressed the need for further practice, exposure, and training using SPSS. They suggested that more training opportunities and

exposure to SPSS package would have been beneficial. These students suggested embedding SPSS demonstrations into lectures, or tutorials to increase students' familiarity with SPSS.

The volunteer students, who study Bachelor in Mathematics, attended training sessions in SPSS, SAS and R. These participants characterised SPSS as an inflexible statistical package that produces voluminous output from individual datasets. Furthermore, they commented that although SAS is more complicated than SPSS, it provides a graphical point-and-click user interface for non-technical users and more advanced options through the SAS programming language. They discussed how in SAS there are more than 300 procedures and each one contains a substantial body of programming and statistical work. Nevertheless, there is no need to use any programming scripting language when working with SPSS since the package does not require from students to learn a whole scripting language. This probably means that SPSS was perceived by students to be the most unlike R and SAS due to their limited programming tools.

These volunteer students proposed increasing the laboratory sessions about SAS to help students to become more competent with using complex aspects of the package. Doing so might help students to build a stronger link between statistical concepts covered in lectures and the technological skills covered during SAS training. A few students raised the complexity of more advanced options provided through SAS and R programming language as being a barrier for students' ability to consolidate their SAS, and R training skills required to operate the SAS and R packages and make them an integral tool used for their statistical investigations.

Almost all participants, who study Bachelor in mathematics, argued that the price of personal license for statistical packages such as SAS and SPSS is a barrier to students. Lack of access to statistical packages may have decreased the effectiveness of training as students have limited opportunities to practice outside of regular training hours provided during their course attendance. Almost all these participants recognised that the free license availability of R package promote students' skills development providing students with valuable practice opportunities outside regular course activities. However, while these participants considered R being accessible in terms of its free license, they expressed dissatisfaction about the accessibility of its user interface and characterized its user interface to be in par with a package like SPSS. Some of the participants, who study Bachelor in psychology, argued that R is hard to learn, however they found programming in SAS harder than in R. They commented that R is very good at data manipulation (merging, recoding, etc.) compared to SAS. SAS was perceived by students to be very good at big data handling for large estimation problems. Moreover, they argued that SAS and R can be used for advanced statistical analyses due to their characteristics such as advanced graphics, while SPSS can be used by students who study social statistics.

The volunteer pre-service secondary mathematics teachers, who attended training session in GeoGebra, seemed to recognise the importance of GeoGebra software for learning introductory statistical concepts and make doing statistical analysis easier when teaching at secondary schools. Similarly, these volunteer pre-service secondary mathematics teachers who participated in the research appeared to recognise that the learning of statistical packages such as SPSS, SAS and R, is important for experts coming from fields that require a strong statistical background to perform advanced statistical analysis on data.

One major concern expressed by some volunteer pre-service secondary mathematics teachers was their difficulties in drawing connections between developing an understanding of statistical concepts and their ability skills. Fifteen participants found that statistics technology skills required for mastering GeoGebra could be used to complement the development of statistical knowledge. Having statistical knowledge and statistics technology skills complementing each other supports Moore's (1997) conjecture that effective learning emerges from the alignment of content, pedagogy and use of technology.

Almost all pre-service secondary mathematics teachers discussed how GeoGebra better support school students' natural urge to explore novel statistical analyses or statistical problems that

were not covered during lectures. Particularly, they argued that GeoGebra supports this enthusiasm by providing opportunities to explore the statistical concepts demonstrated by the applets and figure out how these applets are working. There is perhaps a way to take advantage of students' innate urge to explore GeoGebra applets, by integrating into their lectures more applets created with GeoGebra to demonstrate particular concepts.

Future research along these lines might prove fruitful for promoting the development of statistical learning alongside the development of statistical technological skills when using GeoGebra in teaching and learning Introductory Statistics. Lectures and training methods that promote the development of such skills may prove to be highly beneficial to students during their studies and later in their careers.

The ultimate goal in statistics education is to provide students with the tools to understand statistics. Part of this is being able to use technology to develop understanding, and GeoGebra plays a useful role in that effort to help students take advantage of the technology.

References

- [1] Pratt, D., Davies, N., & Connor, D. (2011). The role of technology in teaching and learning statistics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 97-107). New York: Springer Science+Business Media B.V. 2011.
- [2] Cobb, P., Gravemeijer, K. P. E., Bowers, J. & McClain, K. (1997). *Statistical minitools*. Nashville, TN/Utrecht: Vanderbilt University and Freudenthal Institute/Utrecht University.
- [3] McClain, K. (2008). Dynamic, interactive documents for teaching statistical practice. *International Statistical Review*, 75(3), 295-321.
- [4] Hollylynne, S. L. (2013). Quantitative reasoning in digital world: Laying the pebbles for future research frontiers. Plenary talk at STEM Symposium on Quantitative Reasoning June 1, 2012.
- [5] Henry, M., Girard, J., & Chaput, B. (2008). Modelling and simulations in statistics education. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.) *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey, Mexico: International Commission on Mathematical Instruction and International Association for Statistical Education. Online: www.stat.auckland.ac.nz/~iase/publications
- [6] Abrahamson, D., & Wilensky, U. (2007). Learning axes and bridging tools. *International Journal of Computers for Mathematics Learning*, 12(1), 23-55.
- [7] Engel, J., Sedlmeier, P., & Worn, C. (2008). Modelling scatter plot data and the signal-noise metaphor: Towards statistical literacy for pre-service teachers. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (2008).
- [8] Prodromou, T., & Pratt, D. (2006). The role of causality in the coordination of two perspectives on distribution within a virtual simulation. *Statistics Education Research Journal*, 5(2), 69-88. Online: www.stat.auckland.ac.nz/serj/
- [9] Doer, H., & Pratt, D. (2008). The learning of mathematics and mathematical modelling. In M. K. Heid & G. W. Blume (Eds.), *Research on technology in the teaching and learning of mathematics: Syntheses and perspectives. Mathematics learning, teaching and policy* (Vol. 1, pp. 259-285). Charlotte, NC: Information Age.
- [10] Ben-Zvi, D. (2007). Using wiki to promote collaborative learning in statistics education. *Technology Innovations in Statistics Education*, 1(4). Online: repositories.cdlib.org/uclastat/cts/tise/
- [11] Nolan, D., & Temple Lang, D. (2007). Dynamic, interactive documents for teaching statistical practice. *International Statistical Review*, 75(3), 295-321.
- [12] Pea, R. D. (1987). Cognitive technologies for mathematics education. In A. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 89-122). Hillsdale, NJ: Erlbaum.

- [13] Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning*, 2(1&2), 127-155.
- [14] Hohenwarter, M. (2001). GeoGebra. Available from www.geogebra.com
- [15] Lavicza, Z., Hohenwarter, M., Jones, K., Lu, Y., & Dawes, M. (2010). Establishing a professional development network around dynamic mathematics software in England. *International Journal for Technology in Mathematics Education*, 19(2), 245-252.
- [16] Lee, H. S., Kersaint, G., Harper, S., Driskell, S. O., & Leatharm, K. R. (2012). Teachers' statistical problem solving with dynamic technology: Research results across multiple institutions. *Contemporary Issues in Technology and Teacher Education*, 12(3). Retrieved from <http://www.citejournal.org/vol11/iss3/mathematics/article1.cfm>
- [17] Ainsworth, S. E. (1999). A functional taxonomy of multiple representations. *Computers & Education*, 33, 131-152.
- [18] Finzer, W. (2002). *Fathom dynamic data software* (version 2.1) [Computer software]. Emeryville, CA: Key Curriculum Press.
- [19] Prodromou, T. (2013). Students' emerging expressions of uncertainty while making informal statistical inferences about data. *Proceedings of the joint IASE/IAOS international conference of Statistics Education for progress organised by the International association of Statistics Education (IASE) and the International Association for Official Statistics (IAOS)*, In S. Forbes and B. Phillips (Eds). Macau, China: IASE/IAOS. Online: [http://iase-web.org/documents/papers/sat2013/IASE IAOS 2013 Paper 1.2.4 Prodromou.pdf](http://iase-web.org/documents/papers/sat2013/IASE_IAOS_2013_Paper_1.2.4_Prodromou.pdf)
- [20] Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 65, 123-137.