Journal of Statistical Software

October 2006, Volume 16, Issue 11. http://www.jstatsoft.org/

SOCR: Statistics Online Computational Resource

Ivo D. Dinov
University of California, Los Angeles

Abstract

The need for hands-on computer laboratory experience in undergraduate and graduate statistics education has been firmly established in the past decade. As a result a number of attempts have been undertaken to develop novel approaches for problem-driven statistical thinking, data analysis and result interpretation. In this paper we describe an integrated educational web-based framework for: interactive distribution modeling, virtual online probability experimentation, statistical data analysis, visualization and integration. Following years of experience in statistical teaching at all college levels using established licensed statistical software packages, like STATA, S-PLUS, R, SPSS, SAS, **Systat**, etc., we have attempted to engineer a new statistics education environment, the Statistics Online Computational Resource (**SOCR**). This resource performs many of the standard types of statistical analysis, much like other classical tools. In addition, it is designed in a plug-in object-oriented architecture and is completely platform independent, web-based, interactive, extensible and secure. Over the past 4 years we have tested, fine-tuned and reanalyzed the **SOCR** framework in many of our undergraduate and graduate probability and statistics courses and have evidence that **SOCR** resources build student's intuition and enhance their learning.

Keywords: cooperative learning, collaborative learning, distance education, telelearning, distributed learning environments, improving classroom teaching, interactive learning environments.

1. Introduction

1.1. General

Statistical analyses commonly involve a problem identification and description, data design and acquisition, theoretical model development, manual or automated data analysis and results interpretation (Whitley and Ball 2002a,b,c,d,e). Most of the statistical pedagogical approaches follow a similar design with an emphasis on statistical thinking and practical aspects of data analysis (Lovett and Greenhouse 2000; Taplin 2003). Undergraduate probability and statistics courses are presently either taught with enough rigor, using classical probability theory, or entirely based on empirical observations (Dinov 2006c). In both cases, there are pedagogically valuable reasons for these choices. However, some motivational, descriptive and practical aspects may be significantly downplayed by solely theoretical or entirely empirical instructional approaches. The first setting appears to be mostly adequate for teaching mathematically oriented students. Most of the statistics, applied science and engineering students prefer a more interactive and hands-on instruction to probability and statistics, with arts and humanity students at the far left spectrum. In-class demonstrations, employment of contemporary probability problems and providing computer simulations are thought to enhance the learning process and improve conceptual understanding.

The latest recommendations of many international pedagogical resources in probability and statistics (e.g., Dear *et al.* 2005; Snell *et al.* 2004, American Statistical Association, `http://www.amstat.org/`) suggest that undergraduate students taking probability and statistics courses should be exposed to real-world problems and be given hands-on experiences in generating, collecting and displaying data, as well as trained in model-design, analysis and result interpretation (Hawkins 1997; Teugels 1997; Cox 1998; Taplin 2003). To address these necessities and improve content delivery in undergraduate statistics and probability courses we have built a dynamic collection of interactive online displays, simulations, games, tutorials, presentations, datasets and other resources.

1.2. Background

Several national and international initiatives for overhauling science undergraduate education have recently attracted significant interest from researchers, educators, policy makers and funding agencies. Some of those are the Multimedia Educational Resource for Learning and Online Teaching (Hanley *et al.* 2005, http://www.merlot.org/), the DOE Gateway to Educational Materials (`http://www.thegateway.org/`), Eisenhower National Clearinghouse for Mathematics and Science Education (`http://www.goenc.com/`), Fund for the Improvement of Postsecondary Education (`http://www.ed.gov/programs/fipsecomp/`), Regional Technology in Education Consortium (`http://www.rtec.org/`). The **SOCR** resource is already part of most of these initiatives (`http://www.SOCR.ucla.edu/htmls/SOCR_Recognitions.html`).

In the past several years various groups across the globe have introduced a number of interactive Internet-based statistical textbooks and computational resources. These include: the webbook **Seeing Statistics**, **SurfStat**, the **Goose Statistics Environment**, **StatSoft**, **HyperStat**, **Statistics at Square One**, **WassarStats**, **ResamplingStats**, **WebStat** and **VLPS**. Some of these have general theoretical treatments, whereas some have field-specific flavors including psychology, social sciences, genetics, medical imaging, etc. There are varying amounts of contextual information, degrees of computational capability, interactivity and portability among these tools.

The main problems encountered by statistics instructors, developers and software designers are the variety of hardware architectures that clients (student users) will employ to access these resources, the constantly evolving software languages and operating systems and difficulties associated with updating the resources (including contextual information, pedagogical presentation and graphical appearance) (Inoue *et al.* 2002).

2. The SOCR resource

The **SOCR** resource comprises of a hierarchy of portable online interactive aids for motivating, modernizing and improving the teaching format in college-level probability and statistics courses. More specifically we engineered a number of applets, user interfaces and demonstrations, which are fully accessible over the Internet. The **SOCR** resources allow instructors to supplement methodological course material with hands-on demonstrations, simulations and interactive graphical displays illustrating in a problem-driven manner the presented theoretical and data-analytic concepts. The **SOCR** framework received international attention following a Science magazine review of our probability and statistics resources (Leslie 2003). A key feature of the **SOCR** resources is that they are naturally extensible, as each new instance of a distribution, analysis, game, modeling tool or graph becomes a plug-in that implements one, or several, of our foundational and well-designed interfaces of these types.

The Statistics Online Computational Resource is a collection of Java applets useful for interactive learning and for motivation of various probability and statistics concepts. **SOCR** consists of seven major categories of resources: interactive distribution modeler, virtual experiments, statistical analyses, computer generated games, data modeler, statistical charts and a collection of additional tools.

Introductory probability courses usually begin by referring to various familiar experiments and games (Velleman 2000). Examples of those include rolling dice, dealing cards, tossing coins, playing roulette or crabs, drawing balls from urns, matching hats to people, etc. Despite the fact that most of these games may seem clearly defined, many students experience difficulties addressing specific questions about such experiments. Some of the problems arise from the need for game abstraction, others are due to inability to create a perfect mental representation of the real experiment. To address these issues we designed a virtual experimentation component of the **SOCR** resource (http://www.SOCR.ucla.edu/) that contains a number of games and virtual experiments that are used to build intuition and confidence in understanding such problems. These tools allow students to perform virtual experiments, change parameters of the experiments, observe the outcomes, record frequency distributions, compute sample statistics and compare sample and population characteristics (e.g., means, distribution shapes, etc.) Figure 1, shows a demonstration of this virtual experimentation framework, using a card experiment. This is a dynamic simulation of a random draw of a

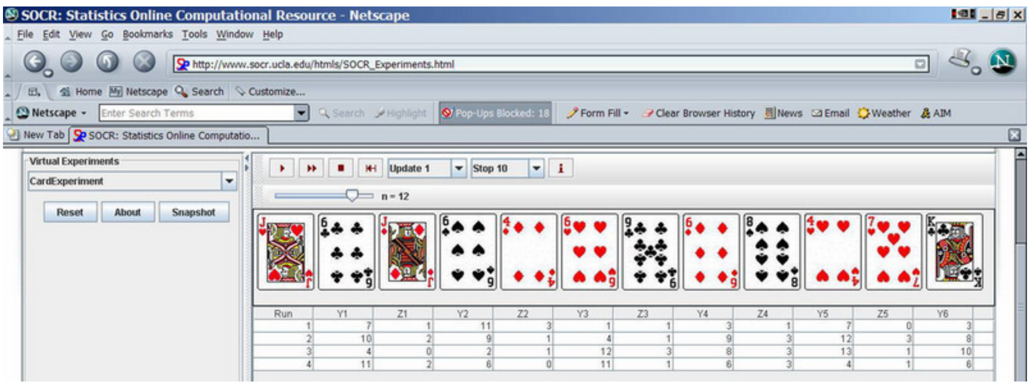

Figure 1: Virtual card experiment: Card drawing simulation and modeling the frequency distributions of outcomes.

hand of cards from a standard 52-card deck.

Students have the ability to set the number of cards that are drawn and control the number of experiments. Summary statistics are computed and displayed following each experiment. Students may draw one hand at a time or run the experiment as a simulation continuously drawing, monitoring the behavior of the sample frequency distribution and comparing it to a specific model distribution.

Another example showing the capabilities of the interactive **SOCR** resources is depicted in Figure 2. This Confidence Intervals Experiment illustrates the relation between the abstract (intuitive) and the procedural (constructive) definitions of confidence intervals. By running this experiment with different values of the confidence level (α) and the sample size (N), students build lasting intuition of the trade-offs between large sample-sizes, high confidence levels and lengths of the resulting statistical intervals. Furthermore, by varying the number of samples drawn (one for each interval) instructors can empirically validate that α*N is the number of intervals that may exclude the estimable parameter of interest, in this case population mean value, $\mu = 0$, (green dots on Figure 2). Previously, static simulations and offline generated graphics have been the best means of such interval consistency validation.

The **SOCR** resource has attracted over 76,000 visitors since 2002. This large number of users translated into over 50,000 active users. These are statistics of users, not hits or page views, and include a single event counting per user per day, no matter how many resources or tools were utilized by the visitor. These numbers also exclude the visits to our educational materials, notes and tutorial provided as additional resources and linked to from **SOCR**. Typical users consisted of UCLA undergraduate students, local researchers and outside students and

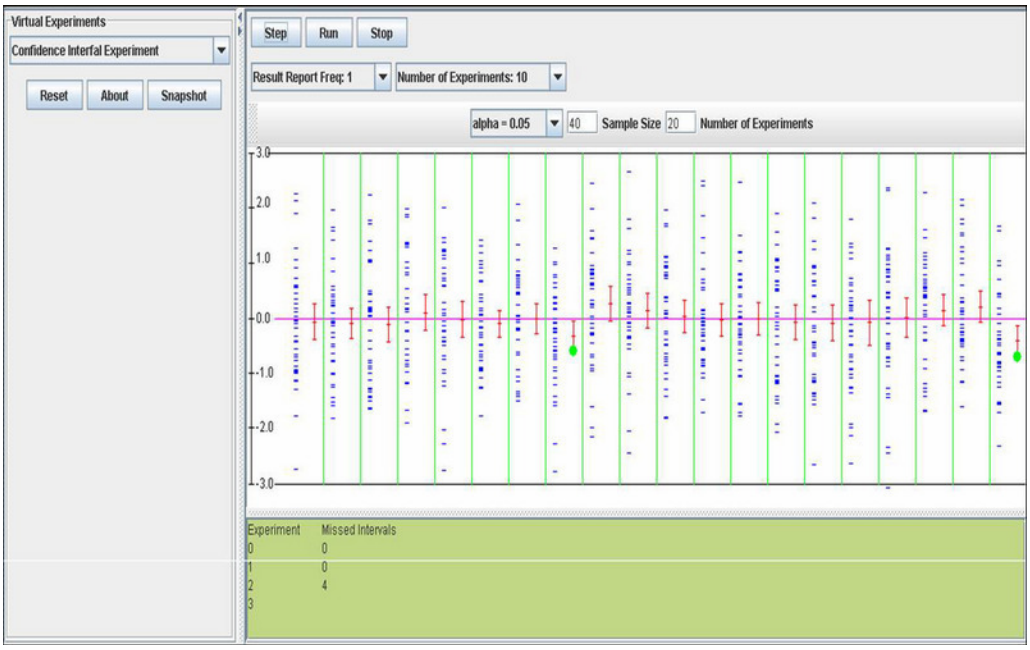

Figure 2: Confidence intervals experiment: A virtual experiment that provides empirical evidence for consistence of the heuristic and the constructive definitions of confidence intervals. The experiment also shows the empirical effects of the sample size and the confidence level on the confidence intervals.

investigators. We have summary statistics of the hourly, daily and monthly **SOCR** usage statistics, as well as browser, operating system, country or origin and type of access to the resources. These are available online at the **SOCR** Geo-Map page and an example is illustrated on Figure 3.

Daily logs of the geographic locations of the last 100 users to the main **SOCR** page (only!) are always dynamically computed and available at the **SOCR** Geo-Map pages, Figure 4, `http://www.SOCR.ucla.edu/SOCR_UserGoogleMap.html`.

A number of license requests, bugs, feature and improvement suggestions were reported by users and reviewers during over the last few years. We have encountered a number of sophisticated expert users, which made some constructive recommendations (e.g., providing numeric high-end 16-decimal precision on probability density function calculations). By cataloging all feedback we are addressing these issues as fast as our resources allow us. For example, we designed a number of JavaScript interfaces that replicate some of our interactive tools,

Reported period	Month Aug 2006				
First visit	01 Aug 2006 - 00:20				
Last visit	31 Aug 2006 - 13:31				
	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic	**3861**	**4716** (1.22 visits/visitor)	**16723** (3.54 pages/visit)	**62093** (13.16 hits/visit)	**2.56 GB** (568.37 KB/visit)
Not viewed traffic			**23191**	**26369**	**557.34 MB**

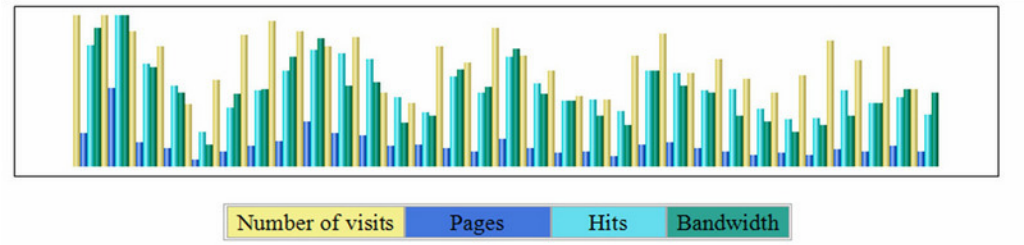

Figure 3: Overall and daily statistics of the **SOCR** resources traffic for the month of August 2006.

Figure 4: National and international **SOCR** resource utilization. This figure illustrates the geographic location of the 100 most recent global **SOCR** visitors on one randomly chosen day (08/12/2006).

however, instead of mouse (imprecise) manipulations and selections we process user control parameters using text fields, to exactly calculate statistics of interest. These tools of primarily computational interest have pedagogical value, as well (e.g., computing of the exact qth quantile for a Normal distribution (Dinov 2006b)).

A suite of distribution modeling aids was designed and implemented in the **SOCR** resource that represents the foundation for sampling/resampling demonstrations, hypothesis tests, statistical inference, model-fitting and critical value estimation. This **SOCR** component allows interactive manipulation of about 40 different families of distributions. Even though the benefit of this tool is limited in real-data analysis situations (by the accuracy of the interactive hand-motion, mouse-precision and screen-resolution) we have received a strong and clearly positive feedback from students, novice users and experts on the intuitive nature, flexible design and the pedagogical potential of this interactive distribution modeling toolbox. Figure 5 illustrates one example where the (standard) Beta (A=0, B=1) distribution is shown for two different pairs of parameters [(α_1, β_1)=(2.2, 10.0) and (α_2, β_2)=(4.9, 2.3)].

Student's attention is brought to the shape of the distribution curve and the effects of the a and b parameters. In addition, instructors may dynamically demonstrate the properties of the distribution model by interactively changing the limits of the region used to calculate the probability of interest.

Many data processing and analysis protocols rely on frequency-based transformations. In a classical setting, data are often times observed in, transformed to, or analyzed in the Fourier

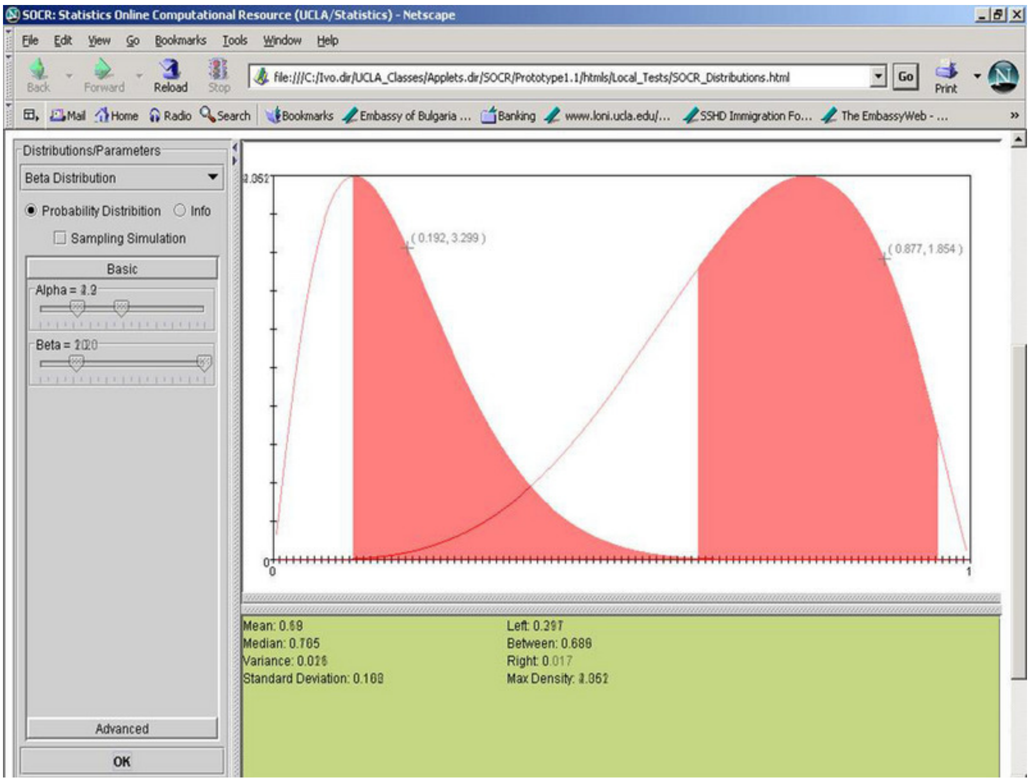

Figure 5: Distribution modeling toolbox: A (standard) Beta distribution, its parameters and shape change and effects of interactive region changes on the probability.

space (Xu and Chan 2002). We have provided a novel design of exploiting multidimensional characteristics of signals in the spatial and frequency domains. This tool is part of the **SOCR** games and allows students to first generate 1D signals (e.g., audio) and then investigate the effects of signal changes onto their Fourier and Wavelet space representations (Dinov *et al.* 2002). Conversely, determining the effects of various frequency, location and intensity-magnitude effects on the special characteristics of the data can also be explored. Figure 6 depicts one such example where wavelet space representation of signals is visually demonstrated. Students are directed to generate manually a periodic 1D signal. The wavelet transformation of the signal is instantly computed and displayed, which allows the student to monitor the effects of altering the frequency, location, magnitude and distribution of the wavelet coefficients of the original signal. A mouse-over event in the bottom wavelet-space panel (yellow) generates the exact magnitude effect of the signal in the spatial domain. All these manipulations are controlled intuitively in real-time by the user using a mouse.

More advanced statistical methods and techniques presented in upper division and graduate level courses may also cause certain level of student discomfort. We consider all interactive and visually appealing hands-on demonstrations as an integral part of the instructional process on such topics as general linear modeling, power analysis, expectation maximization, likelihood ratio inference, mixture modeling, stochastic integration, Brownian motion, Markov Chain Monte Carlo methods, etc. We have begun the high-level design and modeling of Java

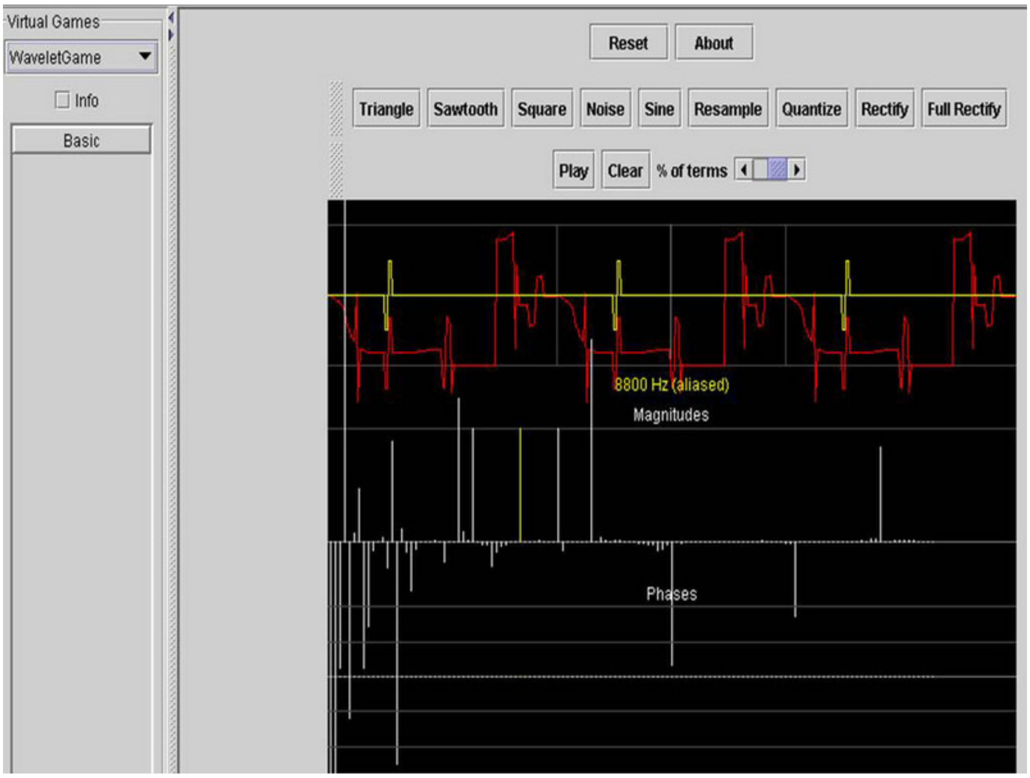

Figure 6: A virtual wavelet game: The interplay between the spatial, frequency, location and signal-intensity parameters and their effects on the data may be subtle and difficult to understand by students. This game demonstrates interactively the specific effects of each of these factors.

applets that illustrate some of these and other powerful statistical techniques. One example that shows both interactive distribution-mixture-modeling and generalized-expectation-maximization was implemented in the setting of 2D point clustering and classification. Figure 7 demonstrates fitting a 3-term Gaussian mixture model in 2D with random isotropic starting kernels (left) and the expectation maximization parameter estimates (final kernel positions, shapes and point classification, right).

Another component of **SOCR** deals with the issue of model fitting. We have designed a framework allowing students to enter data in one of several ways and then attempt to fit a model to their data, see Figure 8.

This approach has no stochastic component built into it, and does not necessarily produce an optimal model. Indeed, one can develop various strategies for analytically optimal computer-based model estimation strategies in some of these situations (cf. previous example, Figure 7). However, in this case we wanted to provide the means for a trial-and-error approach to model fitting - the goal being student intuition development and motivation building. This tool is frequently demonstrated in class to show consistency between theoretical and application driven problem solving, in discussion of the χ^2 goodness-of-fit test, or when validating normality assumptions on data distributions, among others. With this applet instructors can demonstrate the efficacy of statistical testing by sampling/entering data from one distribution (e.g., bimodal 3rd order polynomial) and fitting another distribution model to the data (e.g., Normal distribution or 2nd order polynomial). The resulting χ^2 fit measure, in terms of the corresponding p-value, will be small (rejecting the model) unless the starting sampling distribution and the model distribution are quite similar.

The next component of **SOCR** consists of a cluster of tools for real data analysis. Figure 9 depicts one of the several analyses schemes that we are currently developing. In this example a random sample, from our distribution utility, is generated (dependent values) and a 2-way analysis of variance is performed to identify potential main, simple and interaction effects

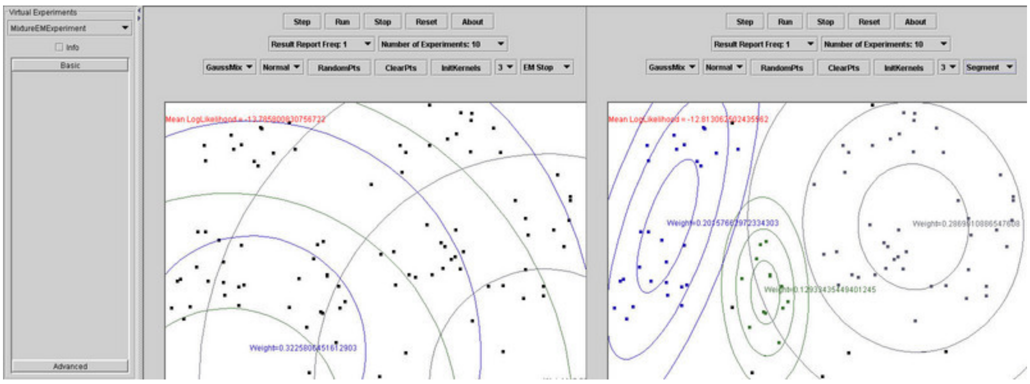

Figure 7: Experiment demonstrating distribution mixture modeling using expectation maximization parameter-estimation: An exploratory tool that demonstrates the effects of the starting mixture of isotropic kernels (location, size) on the final mixture of 3 anisotropic Gaussian models using expectation maximization to estimate the 18 parameters specifying the complete model. The demonstration ends with the classification of all 2D starting points based on which marginal kernel distribution is the most likely candidate that explains their presence in the observed sample (see kernel and point colorings).

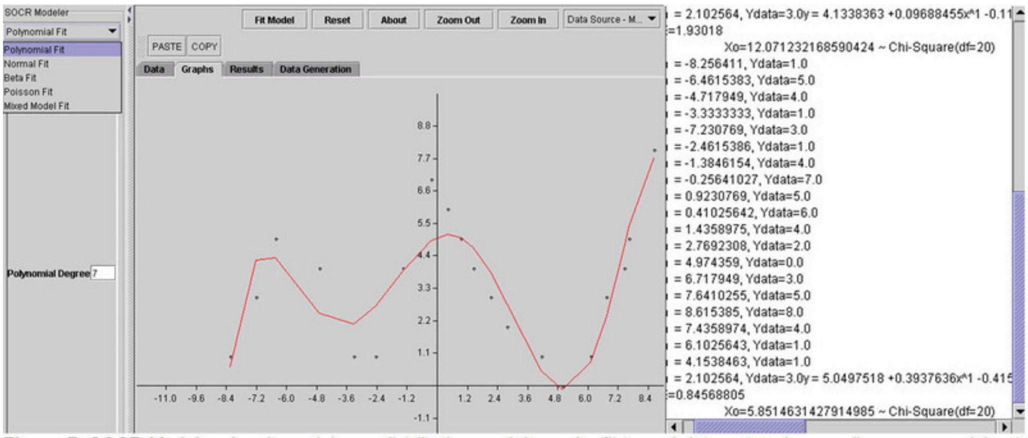

Figure 8: **SOCR** modeler: A polynomial or a distribution model may be fit to real data entered manually as a spreadsheet, read from a file, or virtually generated by mouse clicks. Visual assessment of the model fit may be compared to the analytical model expression and to a statistical assessment of the quality of the fit obtained by using χ^2 goodness-of-fit test. Students may interactively determine the order of the polynomial model, or the distribution parameters of the model, by exploiting different values and observing the visual impact on the model fit.

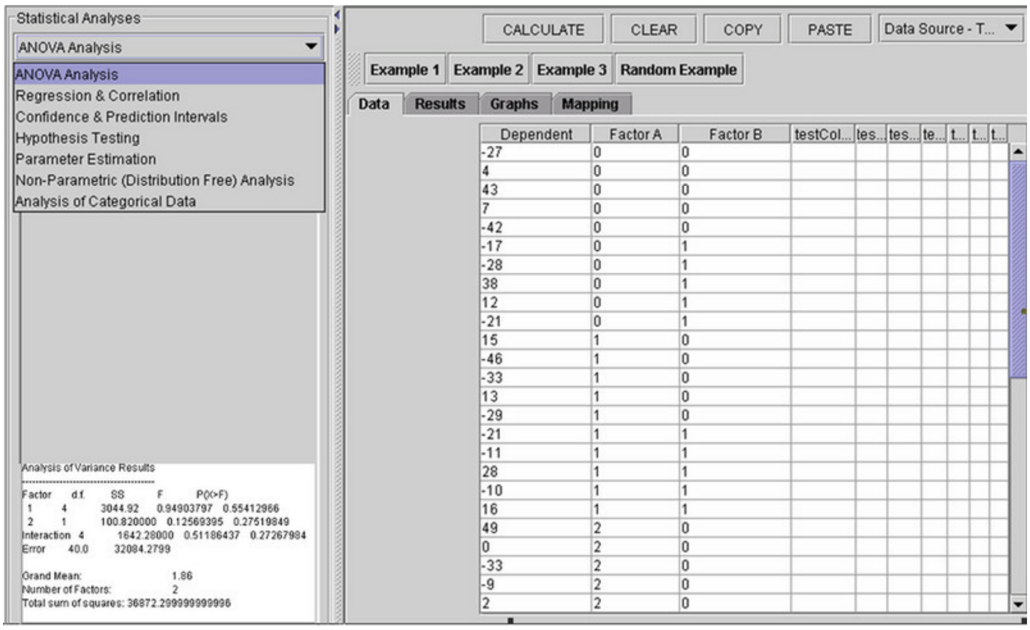

Figure 9: **SOCR** analyses: 2-way analysis of variance (ANOVA) example on randomly generated data.

for the two predictors (factors A and B). Much like in the previous examples students may enter real data into the data table in several different ways and perform a meaningful analysis to determine the strength of the effects of various explanatory factors. The functionality of performing analyses on random data, however, also serves as an important pedagogical instrument to emphasize the consistency between theoretical models and practical analysis protocols, in terms of Type I (false-positive) error, test-statistics and variance decomposition.

SOCR charts is a completely new tool developed as part of the **SOCR** resource in 2006. It is based on **JFreeCharts** (`http://www.jfree.org/`), allows interactive computation of data summary statistics and provides a large number of data plots, charts and diagrams to visually illustrate the characteristics of tabular data. Figure 10, shows one example where we demonstrate a multi-series box-and-whiskers plot along with series-specific summary statistics. **SOCR** charts has a feature that allows the user to look at 3 types of data representations (raw data, data mapping and graphs) and dynamically see the effects of changing the data on the summary statistics and the plot of the data.

The final seventh component of **SOCR** is a collection of additional local and remote resources for probability and statistics education. These include high-precision distribution calculators, online distribution tables, signal plotting and processing tools, other data analysis and demonstration applets. Most of the resources in this category are external and link to outside servers.

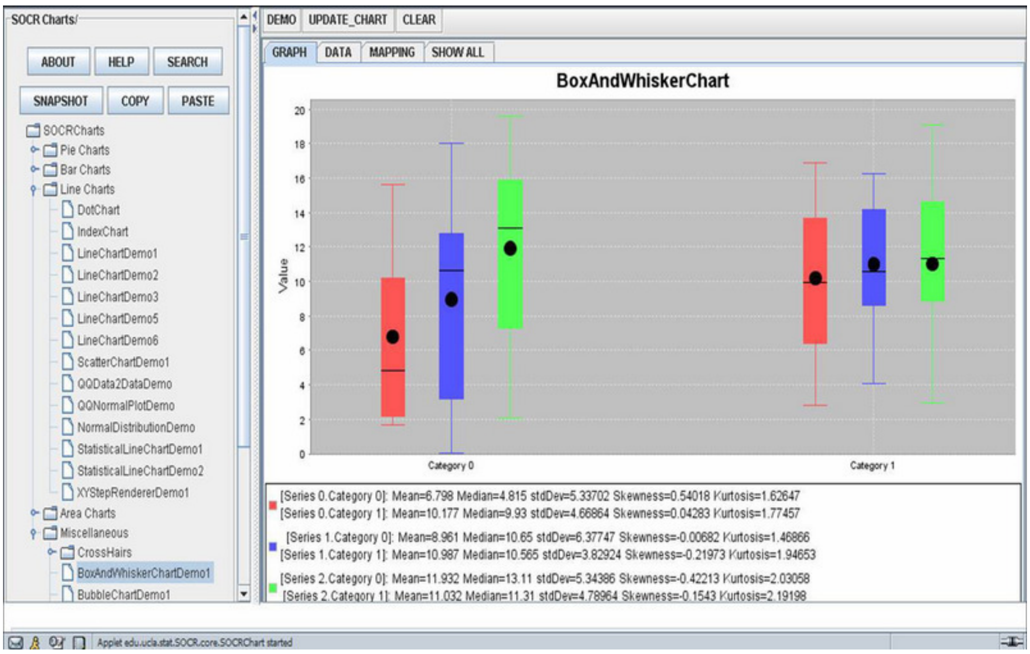

Figure 10: **SOCR** charts: a newly added tool for data summary and graphical data characterization. This is a demonstration of the Box-And-Whisker plot for 2 categories, each with 3 series within it. Summary statistics for all plots are provided on the bottom to complement the qualitative graphical descriptions with their quantitative counterparts.

3. Discussion

The latest recommendations of most international pedagogical resources in probability and statistics (e.g., Dear *et al.* 2005; Snell *et al.* 2004, American Statistical Association, http://www.amstat.org/) suggest that undergraduate students taking probability and statistics courses should be exposed to real-world problems and be given hands-on experiences generating, collecting and displaying data, as well as trained in model-design, analysis and result interpretation (Hawkins 1997; Teugels 1997; Cox 1998; Taplin 2003). With the design of the **SOCR** resource we attempted to do exactly that - develop an extensible and dynamically linked, up-to-date, statistics instructional resource including software applications, instructional materials, virtual experiments, simulations and demonstrations.

3.1. SOCR assessment

In 2005–2006, we introduced over 300 (lower and upper division) undergraduate students to the **SOCR** environment. In three experiments, we compared control groups (using traditional probability and statistics instruction) with treatment groups (**SOCR**-enhanced instruction) in several courses. Our results, (Dinov *et al.* 2006), indicate that the **SOCR** tool utilization in the classroom in probability and statistics classes enhances instruction and student performance. We also observed good outcomes in student satisfaction and use of technology in all three courses. In this study, we also examined the student demographics and their learning styles. In some of the classes, there were statistically significant group differences with respect to the overall quantitative measures of learning (Felder-Solomon Learning Style Index (Felder and Silverman 1998; Felder and Soloman 2003)) and in some of the classes these differences were less clear. We also saw a consistent trend of improvement with time in all **SOCR** treatment groups. Taking in account all assignments and exams for all sections in this study, we observed no examination where the control group (traditional instruction) scored on average higher then the corresponding treatment group (**SOCR**-based instruction). This consistency is indicative of a very strong treatment effect in this study, sign-test p-value$<10^{-5}$.

3.2. Synergies with other similar efforts

A number of outside academic, non-profit and industrial resources were also utilized in the **SOCR** development. We have used ideas, design, tools and models from Elementary Statistics Java Applets (Lewis 2004), **Statlets** (Polhemus 2006), Rice Virtual Labs in Statistics (Lane *et al.* 2005), **StatCrunch** (West 2005), **WebStat** (West 2002), **Statiscope** (Bonnier 2004), **PsychStat** (Stockburger 2005), **BusinessStat** (Arsham 2005), Probability by Surprise (Holmes 2005), Web Interface for Statistics Education (Berger *et al.* 2005), the **CUWU** statistics program (Marden 1997), Probability and Statistics Object Library (Siegrist 2005), **StatLab** (**StatLab** Laboratory for Statistics 2005), Virtual Labs in Probability & Statistics (Siegrist 2006), Java statistics applets at http://www.umd.umich.edu/casl/socsci/econ/StudyAids/JavaStat/applet.htm, **Vestac** (Darius *et al.* 2005), **JSci** (**JSci** Project 2006), **CyberStats** (CyberGnostics Inc. 2003), and many other groups, organizations, student projects, research, clinical and teaching resources. Some of these resources have already been developed with support from NSF/DUE and have significantly impacted the modern pedagogical approaches to teaching undergraduate probability and statistics classes. The Virtual Laboratories in Probability and Statistics (Siegrist 2006) lead by Kyle Siegrist and Dawn

Duehring at the University of Alabama, Huntsville, AL, is one such outstanding example (DUE-9652870/DUE-0089377). This resource provides a number of virtual experiments, games and demonstrations that involve playing cards, tossing dice, flipping coins and other random event trials. Many of those have already been extended and incorporated in one of the **SOCR** components.

Most instructors of contemporary undergraduate science classes utilize the power of the Internet to distribute course related materials on the web. Often these resources are fragmented, not dynamically linked, non-uniform and not very portable. As part of a NSF/DUE CCLI-A&I 9981172 grant (PI: R. Gould) we have developed a web of laboratory assignments for lower division introductory statistics classes for the social sciences, physical sciences, business and economics and biochemistry. These developments (see `http://www.stat.ucla.edu/labs/`) have been well received by students, teaching assistants and instructors. We have also developed, and have been constantly updating, a web of subject-specific class notes resources, homework assignments and study guides (Dinov 2006a) as well as course-specific virtual office hour forums (see `http://www.stat.ucla.edu/forums/`). These instructional materials are extremely useful by themselves for one specific course/section, one discussion topic or one targeted student population. Eventually the **SOCR** resource will serve as an integrated framework of interactive educational tools, online student tutorials, instructional materials, data repository and as an exchange forum. Using web-page access counts, we have established that the current **SOCR** educational materials are utilized repeatedly by the majority of the students enrolled in our undergraduate and graduate classes.

We currently use a 2 GHz dual-processor Mac OS X server to support the **SOCR** resource at `http://www.SOCR.ucla.edu/`. Over 100GB of hard disk space and 1 Giga-bit-per-second Internet connection make the **SOCR** resource robust and responsive. Still, we have a number of developments and improvements to make to complete the **SOCR** resource according to the specified design. Any remote client (user) having Internet connection and a Java-enabled web-browser can access the **SOCR** functionality from any place in the world and at any time of the day. However, there are still great differences in computer hardware, Java-virtual machines, browser versions & manufactures, user settings and firewall/privacy/security preferences that are allowed by modern operating systems, web-browsers and Internet providers. This variation of settings often times causes problems for some users to access the **SOCR** resource. We have tested and verified that the **SOCR** tools are accessible via Java 1.2+ on Mac OS X, IRIX, Solaris and Windows platforms under Netscape 7.2+ and InternetExplorer 6.0+. In addition, **SOCR** is UCLA-signed as a secure applet which may cause problems for some users within restrictive firewall networks. Self-signing was required for interactive data input and output and mouse buffer (cut-and-paste) functionality. We recently obtained an authoritative **SOCR** signing by the major Internet security firm Thawte (`http://www.thawte.com/`). This alleviates many security challenges and browser errors associated with using the **SOCR** resource to access local file system for data input.

The **SOCR** end-user documentation is currently being designed. The **SOCR** developer documentation is already available at the **SOCR** documentation web-page as Java-class hierarchy and as UML diagrams. Currently **SOCR** has a very small kernel and all available features and applications are of plug-in type. This allows direct **SOCR** expansion by outside groups and investigators. The **SOCR** resource is open-source, we share source and binary code and welcome any contributions from the community. In fact, the **SOCR** distributions are currently integrated as the main statistical-distributions package of the **JScience** Project (http://jscience.org/). This is a great recognition for the **SOCR** resource, as **JScience** provides one of the most comprehensive Java libraries for scientific computing and is used in a variety of disciplines (e.g., math, physics, sociology, biology, astronomy, economics, etc.)

We plan to extend the **SOCR** resource to include bi-directional dynamic links between the online computational, demonstrational and analytic resources and a hierarchical HTML-based statistics tutorial, lecture notes, data and problem repositories. Examples of some tools that have achieved certain level of such integration include the outstanding **HyperStat** and NIST Engineering Statistics, among others.

A **SOCR** online discussion forum is also deployed for people to collaborate, share their ideas and possibly form special interest groups, e.g., course- or topic-specific interests. The discussion forum is moderated from time to time for content and suggestions. This helps the online community share their ideas and provide valuable feedback for future developments and improvements to the **SOCR** environment. Specific (meta-) searches based on date, keyword and subject- of the **SOCR** materials/tools/resources/repositories/forums are provided via a site-specific Google search engine. Finally, we have provided easier access points to the **SOCR** resources via dynamic links from http://www.StatisticsResource.org/ and http://www.SOCR.ucla.edu/.

Acknowledgements

Many people have contributed to the **SOCR** development efforts in one form or another over the years. These include, among others, Petros Efstathopoulos, Robert Gould, Guogang Hu, Jianming Hu, Jason Landerman, Fotios Konstantinidios, Hui Wang, Donald Ylvisaker, Dushyanth Krishnamurthy, Jeff Ma, Annie Che, Arthur Toga, Jenny Cui.

A number of outside academic, non-profit and industrial resource were also utilized in the **SOCR** development. We have used ideas, design, resources and models from Elementary Statistics Java Applets, **Statlets**, Rice Virtual Lab in Statistics, **WebStat**, **Statiscope**, **PsychStat**, **BisinessStat**, Probability by Surprise, Web Interface for Statistics Education, **CUWU Stats**, **PSOL**, **StatLab**, Virtual Labs in Probability & Statistics, JavaStat, **Vistac**, **CyberStat**, and many other groups, organizations, student projects, research, clinical and teaching resources.

In addition, the **SOCR** project was supported in part by the following State of California and Federal grants UCLA OID IIP (IIP0318), NIH/NCRR (P41 RR13642), NCBC (U54 RR021813) and NSF CCLI-EMD (044299).

References

Arsham H (2005). "**BusinessStat** – A Collection of JavaScript E-labs Learning Objects." URL http://home.ubalt.edu/ntsbarsh/zero/scientificCal.htm.

Berger D, Berger D, Aberson C, Healy M, Romero V, Saw A, Sosa G, Emerson E, Huntley J, Blanchard A, Kyle D, Prull M, Light L, Nardi P, Pezdek K, Smiley P, Sullivan J (2005). "Web Interface for Statistics Education." URL http://wise.cgu.edu/.

Bonnier M (2004). "**Statiscope**." URL http://www.df.lth.se/~mikaelb/statiscope/statiscope.shtml.

Cox DR (1998). "Some Remarks on Statistical Education." *Journal of the Royal Statistical Society D*, **47**(1), 211–213.

CyberGnostics Inc (2003). "**CyberStats**." URL http://statistics.cyberk.com/splash/.

Darius P, Beirlant J, Lesaffre E, Loosveldt G, Onghena P, Vandebroek M, Vermeire L, Vuylsteke M, Michiels S, Raeymaekers B, Moreau W, Solomin A, Wang R, Van den Noortgaete B (2005). "**Vestac** – Java Applets for Visualization of Statistical Concepts." URL http://www.kuleuven.ac.be/ucs/java/.

Dear K, Smith R, Coombes J, Brennan R (2005). "**Surfstat.australia**: An Online Text in Introductory Statistics." URL http://www.anu.edu.au/nceph/surfstat/surfstat-home/surfstat.html.

Dinov ID (2006a). "Ivo Dinov's Probability and Statistics Class Notes." URL http://www.stat.ucla.edu/~dinov/courses_students.html.

Dinov ID (2006b). "**SOCR** Exact JavaScript Calculators." URL http://www.SOCR.ucla.edu/Applets.dir/Normal_T_Chi2_F_Tables.htm.

Dinov ID (2006c). "**SOCR**: Statistics Online Computational Resource: SOCR.ucla.edu." *ASA Statistical Computing & Graphics Newsletter*, **17**(1), 11–15.

Dinov ID, Mega MS, Thompson PM, Woods RP, Sumners DWL, Sowell EL, Toga AW (2002). "Quantitative Comparison and Analysis of Brain Image Registration Using Frequency-Adaptive Wavelet Shrinkage." *IEEE Transactions on Information Technology in Biomedicine*, **6**(1), 73–85.

Dinov ID, Sanchez J, Christou N (2006). "Pedagogical Utilization and Assessment of the Statistic Online Computational Resource in Introductory Probability and Statistics Courses." *Journal of Computers & Education*, **in press**.

Felder R, Silverman L (1998). "Learning and Teaching Styles in Engineering Education." *Engineering Education*, **78**, 674–681.

Felder R, Soloman B (2003). "Index of Learning Styles Questionnaire." URL http://www.engr.ncsu.edu/learningstyles/ilsweb.html.

Hanley G, Reisman S, Swift C, Carey T, Lalanne A, Leal P, Kasza M, Worthington A, Sperling BB (2005). "Multimedia Educational Resource for Learning and Online Teaching." URL http://www.merlot.org/.

Hawkins A (1997). "Discussion: Forward to Basics! A Personal View of Developments in Statistical Education." *International Statistical Review*, **65**(3), 280–287.

Holmes S (2005). "Probability by Surprise." URL http://www-stat.stanford.edu/~susan/surprise/.

Inoue T, Asahi Y, Yadohisa H, Yamamoto Y (2002). "A Statistical Data Representation System on the Web." *Computational Statistics*, **17**(3), 367–378.

JSci Project (2006). "**JSci** – A Science API for Java." URL http://jsci.sourceforge.net/.

Lane D, Scott D, Benway J, Lu J, Tang Z, Shea A, Quinones M, Baggerly K, Austin J, Swartz M, Swartz R (2005). "Rice Virtual Lab in Statistics." URL http://www.ruf.rice.edu/~lane/rvls.html.

Leslie L (2003). "Statistics Starter Kit." *Science*, **302**(5), 1635.

Lewis B (2004). "Elementary Statistical Java Applets and Tools." URL http://intrepid.mcs.kent.edu/.

Lovett MC, Greenhouse JB (2000). "Applying Cognitive Theory to Statistics Instruction." *The American Statistician*, **54**(3), 196–206.

Marden J (1997). "**CUWU** – The CUWU Statistics Program." URL http://www.stat.uiuc.edu/courses/stat100/cuwu/.

StatLab Laboratory for Statistics (2005). "**StatLab**." URL http://statlab.fon.bg.ac.yu/eng/eng/apletieng/resources/resources4.html.

Polhemus NW (2006). "**Statlets**." URL http://www.statlets.com/.

Siegrist K (2005). "Probability and Statistics Object Library." URL http://www.math.uah.edu/stat/objects/.

Siegrist K (2006). "Virtual Laboratories in Probability and Statistics." URL http://www.math.uah.edu/stat/.

Snell JL, Doyle P, Garfield J, Moore T, Peterson B, Shah N (2004). "Chance Project." URL http://www.dartmouth.edu/~chance/.

Stockburger DW (2005). "**PsychStat** – Psychological Statistics at Missouri State University." URL http://www.psychstat.missouristate.edu/.

Taplin R (2003). "Teaching Statistical Consulting Before Statistical Methodology." *Australian & New Zealand Journal of Statistics*, **45**(2), 141–152.

Teugels JL (1997). "Discussion: Forward to Basics! A Personal View of Developments in Statistical Education." *International Statistical Review*, **65**(3), 287–288.

Velleman PF (2000). "Design Principles for Technology-based Statistics Education." *Metrika*, **51**(1), 91–104.

West W (2002). "**WebStat** 3.0 – Statistical Software for Data Analysis on the Web." URL http://dostat.stat.sc.edu/webstat/3.0/oldindex.html.

West W (2005). "**StatCrunch** – Statistical Software for Data Analysis on the Web." URL http://www.statcrunch.com/.

Whitley E, Ball J (2002a). "Statistics Review 1: Presenting and Summarising Data." *Critical Care*, **6**(1), 66–71.

Whitley E, Ball J (2002b). "Statistics Review 2: Samples and Populations." *Critical Care*, **6**(2), 143–148.

Whitley E, Ball J (2002c). "Statistics Review 3: Hypothesis Testing and P Values." *Critical Care*, **6**(3), 222–225.

Whitley E, Ball J (2002d). "Statistics Review 5: Comparison of Means." *Critical Care*, **6**(5), 424–428.

Whitley E, Ball J (2002e). "Statistics Review 6: Nonparametric Methods." *Critical Care*, **6**(6), 509–513.

Xu Z, Chan A (2002). "Encoding with Frames in MRI and Analysis of the Signal-to-Noise Ratio." *IEEE Transactions on Medical Imaging*, **21**(4), 332–342.

Affiliation:

Ivo D. Dinov
The **SOCR** Resource
Department of Statistics
8125 Mathematical Science Bldg.
University of California, Los Angeles
Los Angeles, CA 90095-1554, United States of America
Tel. +1/31/825-8430
Fax: +1/31/206-5658
E-mail: dinov@stat.ucla.edu
URL: http://www.SOCR.ucla.edu/

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/

Volume 16, Issue 11 *Submitted:* 2005-07-01
October 2006 *Accepted:* 2006-10-01