Origins of the NIST/SEMATECH e-Handbook of Statistical Methods in the Work of Mary Natrella

Carroll Croarkin and Will Guthrie, Statistical Engineering Division, NIST

The NIST/SEMATECH e-Handbook of Statistical Methods[1], is a Web-based book whose goal is to help scientists and engineers incorporate statistical methods into their work as efficiently as possible. Ideally it will serve as a reference that will help scientists and engineers design their own experiments and carry out the appropriate analyses when a statistician is not available to help. It is also hoped that it will serve as a useful educational tool that will help users of statistical methods and consumers of statistical information better understand statistical procedures and their underlying assumptions and more clearly interpret scientific and engineering results stated in statistical terms.

The project began with a request from Patrick Spagon of the Statistical Methods Group of SEMATECH, a consortium of major U.S. semiconductor manufacturers, to update the National Bureau of Standards[2] *Handbook 91*, *Experimental Statistics*. Written by Mary Natrella of the NBS Statistical Engineering Laboratory and published in 1963, *Handbook 91*, was a best-selling NBS publication for many years. Engineers and scientists in a variety of fields appreciated it because of its problem-oriented approach to statistics and its detailed examples. In fact, Pat Spagon had been using it extensively in his teaching at SEMATECH and was especially interested in replacing the existing examples in the Handbook with examples related to the semiconductor community's needs.

To understand the appeal of *Handbook 91*, it is helpful to know something of Mary Natrella and the impact of her career at NBS. Mary was an expert on the application of modern statistical techniques in physical science experimentation and engineering testing and worked for thirty-six years as a statistician at the National Bureau of Standards before retiring in 1986. The multi-faceted nature of her career involved heavy emphasis on teaching statistical methods to NBS scientists and consulting on applications of statistical methods in the physical sciences and engineering. Early in her career, she participated on the team that produced the initial statistical sampling standards document (MIL-STD-105A) that came to be used throughout the world. This background made her an ideal candidate for organizing and producing a handbook for presenting statistical methods.

Mary was a fellow of the American Society for Test Materials (ASTM) in recognition of her work on ASTM E-11, the sub-committee responsible for the development of statistical standards. More recently, the Quality and Productivity Section of the American Statistical Association established the endowed Mary G. Natrella Scholarship

[1] http://www.nist.gov/stat.handbook/

[2] now the National Institute of Standards and Technology (NIST), an agency of the U.S. Department of Commerce whose mission is to strengthen the U.S. economy and improve the quality of life by working with industry to develop and apply technology, measurements, and standards.

which is presented annually to two graduate students interested in pursuing quality and productivity in industry using statistical methods.

Mary felt very strongly that examples of statistical procedures in *Handbook 91* should also accompanied by fill-in-the-blank worksheets with sample data, allowing the reader to quickly and easily duplicate the calculations with his or her own data. Mary's *Handbook* was replete with such examples By the 1990's, however, when SEMATECH and NIST began their joint collaboration, the emphasis on hand calculations was too dated to be practical and many modern statistical methods were missing from the text.

As a result, a joint NIST/SEMATECH project team was assembled to explore the feasability of producing a new Handbook, develop a formal project proposal and carry the project out. Paul Tobias headed the SEMATECH team which included Patrick Spagon, Barry Hembree, Jack Prins and Chelli Zey. Carroll Croarkin headed the NIST team which included James Filliben, Willliam Guthrie and Alan Heckert. The first meeting was hosted by Dominic Vecchia of the Statistical Engineering Division office in Boulder, Colorado.

With the rapid growth of the Internet at that time, the project quickly evolved from the publication of a new edition of a traditional book to development of an online handbook for distribution via the World-Wide Web. The advantages of Web-distribution with easy access by users all over the world and the opportunity to create an easily expandable resource were obvious.

However, Mary's original vision of a reference book with easily duplicated examples remained a guiding principle for the new handbook. The idea of abandoning hand-worked examples in favor of software for illustrating statistical methods in the new handbook eventually led to a decision to integrate the necessary software directly into the Handbook. The freely available statistical package DATAPLOT (originally developed by James Filliben and maintained by NIST) was expanded to meet these needs, and an interface was developed to allow the reader to run this software from within the handbook. Thus, the team was able to perpetuate Mary's vision and maintain a practical, problem-oriented approach to statistics by illustrating the methods with worked examples using data from the NIST and SEMATECH laboratories.