Introduction:
How to Read this SOAR

Jorge Luis Romeu
IIT Research Institute
May 28, 1999

Sometimes, materials engineers have problems understanding the philosophy behind the statistical procedures they need to apply, when analyzing materials data. This is not surprising for us. We have found, after twenty-five years of teaching statistics at the undergraduate, graduate and professional training levels, that the majority of professionals obtain their statistical training during their undergraduate –since most do not proceed to graduate school. In addition, we also found (i) that undergraduate statistics courses are usually limited to, at most, three; (ii) that these courses are rather theoretically oriented and (iii) that they discuss way too many statistical techniques in too short a time. None of these situations help students to mature or obtain a full appreciation of statistical concepts and uses.

As a consequence, statistics is usually one of the most disliked subjects in student life. Instructors are not able to motivate many students and thus, cannot teach them the subject matter fully. Since most students never take further formal statistics education, many end up thinking it is an activity where one crunches some numbers, via some time consuming formulas, compares the results to an incomprehensible table and obtains some probability statements on which they do not place much faith.

This serious problem has been getting much attention and efforts for its correction (see references 1, 2, 3, 4 and suggested readings at the end of the chapter) have already started. For, students become engineers and need to work with data, to perform statistical analyses and to use their results in real life work. Then, they discover their lack of proper statistical training and realize that they must now get these on the job. To aid engineers to obtain such on-the-job training, many guides and handbooks have been developed [e.g. 5, 6, 7]. But these are usually right to the nuts and bolts, full of statistical formulas and detailed step-by-step procedures and quite crowded with many different topics. Therefore, in spite of their good quality, these handbooks, by themselves, may not help the engineers much more than college courses helped the students -and for the same reasons above mentioned.

The objective of the present SOAR is, precisely, to fill such a gap. For, many engineers and practitioners, who deal with materials data analyses, may still have a strong need to understand the statistical methods in order to implement them more efficiently and to interpret their results more accurately. Hence, the overall goal of this SOAR is to provide both, the basic understanding and the statistical thinking or philosophy, behind the procedures in the materials data analysis handbooks [6 and 7]. Such better understanding helps to raise the overall level of the data analysis activity.

This SOAR originates in the warm response received to a series of statistics articles published in the MaterialEASE insert of the AMPTIAC Newsletter [1 through 4]. These articles discussed statistical thinking rather than the nuts and bolts of the materials handbook procedures. Expanding and building on this experience, this SOAR will provide hands-on experience, by going over completely developed and commented practical examples of materials data analyses. In addition, the many SOAR problem discussions, observations and considerations, combined with data explanations and problem extensions, will provide insight on the application of these methods.

The technical level required for reading this SOAR is introductory. We will assume the reader meets two basic prerequisites. First, the reader should have some exposure to introductory statistics, via a general course (usually taken sometime back) or by having worked in materials data analysis. Then, the reader should also have some background, by formal training or by hands-on experience, in material science. Both of these prerequisites intend provide some exposure about materials analysis, which is all that is needed to successfully read this SOAR.

There are eight chapters here. Chapter one explains, via some discussions and practical examples, why the materials engineers should read this book and why it is relevant to them. Chapters two through five are based on the four referenced AMPTIAC newsletter MaterialEASEs [1 through 4]. They deal, respectively, with data and its pedigree, distributions and parameters, estimation and testing and regression and analysis of variance. These chapters present practical, illustrative examples that enhance the commentaries and explanations.

In addition, chapters six and seven discuss and develop, respectively, case studies of materials data analyses using methods for regression and ANOVA from MIL HDBKs 5 and 17. These two chapters develop specific expertise in these two areas and constitute the core of this SOAR while the first five chapters intend to take the reader up to speed in the use of statistics, to help them better understand the latter two. A final chapter summarizes the methods and results presented in this SOAR and discusses possible future directions. A diskette with the data sets analyzed is included, for readers to practice.

This SOAR does not intend to be another statistics textbook or to compete with other guides or MIL Handbooks 5 and 17. Instead, it intends to complement them, by helping practicing engineers understand and make a better use of these excellent materials. For, textbooks are not targeted for the specialized materials data analysis engineering work. Hence, the methods required are sparsely distributed among different books and chapters and at different technical levels. On the other hand, the methods in MIL HDBKs 5 and 17 are presented as cut and dry step-by-step procedures, without great explanations or discussions about them.

Some handbook examples will be reworked in a more threaded way (as in case studies). They will also include discussions and extensions, with frequent references made to the handbooks as well as to five selected statistics textbooks [8, 9, 10, 11 and 12]. The selection of these reference materials was based on two reasons. First, they are excellent

and very accessible textbooks and reports. Then, we have had ample experience teaching with them, both general and engineering statistics courses, and our students have always found them useful and easy to read.

All concepts are introduced through three types of illustrative and practical examples. First, some materials data sets have been specifically obtained for this purpose. Then, some simulation examples were created to illustrate specific statistical problems in chapters two through five. Finally, some examples for chapters 6 and 7 are taken directly from [5, 6 and 7] and are reworked here. In their original source, they were either only partially developed or were developed differently. They are specifically used here to provide comparisons that yield important insight on the correct use of statistical methods.

The audience for this SOAR includes four large groups of people that deal with materials data analysis. First, it includes the data analysts, who need to correctly implement statistical procedures on materials data. Secondly, it includes the designers, who need to obtain accurate allowables for their designs. This SOAR will help them both to sharpen their skills and their understanding of how to better use these procedures. Thirdly, it includes the data gatherers, who provide the raw material for statistical analysis. By improving their understanding of the analysis process, they can gather better data and help the analysts to do better statistics. For, statistical analysis can be just as good or as bad as the data it deals with. Finally, the managers also have a big stake in this process. For, without their understanding of, and the support for, the important, costly and time consuming data gathering, processing and analysis activity, there is very little that any of the other above mentioned groups can do.

Three, non-necessarily excluding results are expected from reading this SOAR. First, readers will gain a firm understanding of, and an appreciation for, materials data statistical analysis. This is especially useful for data gatherers and managers, who are indirectly but relevantly related to this analysis process. Secondly, the analysts will better understand the statistical thinking underlying the procedures in MIL HDBKs 5 and 17. This will allow them to obtain stronger practical analysis results. Finally, all readers will also develop better hands-on skills, when implementing the statistical procedures in materials data analysis.

Below, we present a comparative table that summarizes all the objectives of this SOAR, compares them to those of handbooks and textbooks and frames this discussion.

From the table below we can read how courses and textbooks provide the most complete treatment of statistical theory and formulas. If readers have the time and facilities, this remains their best option. However, books are too general and courses would take too long to follow this path if one wants to obtain a working knowledge of the subject in a reasonably short period of time.

Aspect	Characteristics and SOAR	Comparisons for AMPTIAC Handbook	SOAR Textbooks
Theory	Philosophy	Basics	Complete
Formulas	Important ones	Most formulas	All formulas
Examples	Threaded and related	Maze-like examples	Unrelated to field
Computer Work	Specific examples	Some examples	Few examples
Work Focus	Concentrated	Dispersed	General Interests
Specialization	Very specialized	Specialized	No specialization
Independence	Some	Most	Complete
Material Level	Introductory	Specialized	Advanced/General
Reading Sequence	First or Companion	Second one	Last one
Overall Effect	Understanding	Hands-on work	General Knowledge
Pre-Requisites:	exposure (one semester) materials engineering	of basic statistics background	
Ideal Sequence:	junior materials engineer study of the SOAR Study of the Handbook Then take a specialized	or technician background material and exercises materials and exercises. statistics course in college	

We also see that the handbooks, on the other hand, rapidly present the basics of the specialized theory, providing step-by-step procedures for learning the mechanics of the materials analysis work through a maze of examples. This constituted a great improvement. However, they leave out many explanations about the philosophy or use of the statistical concepts they describe. And this hurts the learning process of the isolated reader. In addition, their extensive material is widely dispersed throughout the handbook, which makes them difficult to follow and read. This SOAR, in turn, provides a starting point for the beginner who wants to study these topics and a bridge between the more experienced practitioner and these handbooks and materials.

A suggested reading sequence of these materials is as follows. Engineers or technicians with little or no experience in materials data analysis, but with the stated prerequisites, would start by reading the SOAR and working out its examples. Then, as they become acquainted with the methods, they can go on to read the handbook procedures [6 and 7] using this SOAR as a "study guide". Finally, after acquiring mastery and understanding of the data analysis methods, readers could continue on to specialized textbooks and course offerings.

The above roadmap clearly shows the need for a guide that helps the practitioner get through this maze of knowledge and provides the missing explanations and insight to understand them.

This SOAR provides just that.

Suggested Further Readings

How University-Industry Partnership Ought to Work. By George M. Low. RPI Alumnae Magazine. Summer 1983
Academe Meets Industry: Charting the Bottom Line. By Elise Hanckock. RPI Alumnae Magazine. August 1983
Providing Effective Training in Industry: An Experience. By T. P. Callahan and H. W. Altland. ASA 1995 Proceedings on Statistics Education.
Non Mathematical Statistics: A New Direction for the Discipline. Special Section. Undergraduate Education. The American Statistician V. 53 No. 1 (1999).