

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

USDA National Wildlife Research Center - Staff
Publications

U.S. Department of Agriculture: Animal and
Plant Health Inspection Service

November 1984

Statistical Consulting in a Geographically Dispersed Organization

Richard M. Engeman

Denver Wildlife Research Center, U.S. Fish and Wildlife Service, s_r100@yahoo.com

William E. Dusenberry

Denver Wildlife Research Center, U.S. Fish and Wildlife Service

Follow this and additional works at: https://digitalcommons.unl.edu/icwdm_usdanwrc



Part of the [Environmental Sciences Commons](#)

Engeman, Richard M. and Dusenberry, William E., "Statistical Consulting in a Geographically Dispersed Organization" (1984). *USDA National Wildlife Research Center - Staff Publications*. 163.
https://digitalcommons.unl.edu/icwdm_usdanwrc/163

This Article is brought to you for free and open access by the U.S. Department of Agriculture: Animal and Plant Health Inspection Service at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in USDA National Wildlife Research Center - Staff Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

In this section, *The American Statistician* publishes articles and notes of interest to teachers of the first mathematical statistics course and of applied statistics courses. To be suitable for this section, articles

and notes should be useful to a substantial number of teachers of such a course or should have the potential for fundamentally affecting the way in which the course is taught.

Statistical Consulting in a Geographically Dispersed Organization

RICHARD M. ENGEMAN and WILLIAM E. DUSENBERRY*

Tis distance lends enchantment to the view

Thomas Campbell, 1799

There has been much discussion in the recent statistical literature concerning the optimal manner for a consulting statistician to operate within a larger organization. Marquardt (1979) described the "total involvement" approach, involving a statistician from conceptualization of a project to the final report; Finney (1982) discussed the questions a statistician should ask for effective consulting; and Bishop, Peterson, and Tragser (1982) argued that a statistician's most substantive contributions are made during the planning stages of a study. McCulloch et al. (1982) described a program for training statistical consultants, and Hunter (1981) discussed the roles a statistician may assume.

A statistician consulting in a geographically dispersed organization will encounter difficulties similar to those of a statistician consulting in-house. The typical obstacles encountered through in-house consulting are greatly exaggerated, however, when consulting with clients at distant sites. Similarly, the skills needed to be an effective consultant at a distance are generally the same as those required for in-house consultation, but some must be enhanced for effective consultation at a distance. In this article we present, from our experience, some of the problems that become exaggerated when consulting at a distance, some of the skills that must also be exaggerated to produce effective consulting in this situation, some features we have found useful in long-distance consultation, and some recommendations for the classroom, which may help prepare students for consulting by telephone.

As with in-house consulting, visits to experimental sites are an important aspect for effective statistical consultation (e.g., Hahn 1984). A statistician consulting with distant field sites has a different involvement in a site visit, however, than does a statistician consulting in-house. Hahn stated that the major mode of communication between experimenter and statistician should be face to face. A statistician

consulting in-house is available for face-to-face communication as needed by the experimenter, in addition to visiting the experimental site. For a statistician consulting with distant sites, a site visit frequently represents the *only* opportunity for face-to-face consultation across a span of months or years. Thus a visit to a distant site is a "temporary" resource, and the statistician must expect to have his or her talents exploited to the fullest. Consulting relationships must be established and developed in a relatively short span of time; current projects must be reviewed and future projects anticipated; and the statistician must learn the purposes and constraints under which the field station is operating. An in-house site visit generally does not need to accomplish as much. When consulting with an in-house experimental site, if something more needs to be done, the statistician and experimenter can always meet again later. When consulting with a distant site, however, the statistician will have to accomplish anything left undone over the telephone, since another visit may not be feasible in the near future. Therefore, a statistician visiting a distant site must expect much longer and more intensive consulting sessions than would be usual for in-house visits. Such visits to distant field sites usually require far more mental and physical stamina than would be expected in-house, and a statistician unwilling to meet the demands will not produce a successful site visit.

Ideally, a statistician should be able to make a site visit at the start of every major project. This is rarely accomplished, even for in-house situations; but visits to distant projects are accomplished far less frequently than nearby or in-house projects. In practice, it may be fortunate if a statistician ever visits a particular field site, let alone at the start of each project. Consequently, most statistical consulting in a widely dispersed organization is done by telephone. A statistician in a central location must therefore sharpen the skills that allow consultations over the telephone to be as effective as possible.

When consulting over the telephone, the statistician must not only receive and assimilate complicated information but be able to convey complex ideas in a manner that is useful to the person in the field. In the absence of face-to-face communication, the difficulty of this task is greatly magnified. Possibly the most essential element for effective consultation by telephone is the ability to understand the situation quickly (objectives, study design, data structure, etc.). The statistician needs to ask precise questions that are

*Richard Engeman is a Statistician, and William Dusenberry is a Statistician and Chief of the Section of Technical Services, both at the Denver Wildlife Research Center, U.S. Fish and Wildlife Service, Bldg. 16, Denver Federal Center, Denver, CO 80225. The authors thank Dave Otis, John Oldemeyer, and anonymous referees for their helpful reviews.

likely to produce the required information without visual assistance. A review of a preliminary study plan, if one is available, also promotes quick understanding. Conversely, statistical concepts and rationale must be explained over the phone so that they make sense and are educational to a nonexpert. Designs, analyses, and so forth must be described in a simple, easy-to-follow fashion so that the information described over the telephone is in fact carried out in the field. Because there is no visual contact during a telephone consultation, the accurate receipt and transmission of information and ideas becomes more difficult. We frequently find it useful to diagram the information being sent or received and to have the client do the same. If each party iteratively describes and questions their version of the diagram over the telephone, then misunderstandings are more easily located and the transmission of accurate information is facilitated.

As with many consulting situations, telephone consultation frequently necessitates follow-up work, thus requiring the statistician to be strongly motivated. The amount of follow-up work from a telephone consultation is commonly greater than in an ordinary consultation, since the statistician's resources during the consultation are frequently limited by the length of the phone cord. When the client is at a distance, it is far easier to procrastinate than when the client's physical presence in-house helps to spur the statistician to action. The statistician must, however, give the client in the field priority equal to those who are in-house. If prompt follow-up is not forthcoming, the relevance of the statistician's efforts is reduced in the investigator's mind and the urgency of a field project may override waiting for statistical input. To avoid misinterpretations, the substance of phone consultations should often be documented and mailed to the investigator. This provides a record of the information furnished to the statistician, its interpretation, the statistician's recommendations, and what responsibilities each party has agreed to assume.

When working with field station staff in a widely distributed organization, the consultant should always be aware of such possible influences as culture, politics, customs, and languages that are different from those of the central location. An in-house statistician generally does not have to take these influences into consideration, but a statistician consulting to distant sites must or else risk being misinterpreted or giving advice that cannot or will not be carried out in the field. Such a situation would cause the statistician to lose credibility and cause field studies to suffer from lack of statistical input. We have learned to be cognizant of these considerations while consulting for 26 field stations throughout the United States and in four developing countries (on more than one occasion, we have consulted in Spanish). Understanding the role of cultural influences (some of which are greater barriers than language) enhances the statistician's ability to provide credible advice to distant sites.

The availability of quick, interactive computing is a boon for all statisticians, but particularly so for the telephone statistician. Results can be returned quickly, since necessary computing can frequently be performed during a phone conversation or very soon thereafter. Having such computing

capability available allows the statistician to confirm swiftly some computational efforts made at the field station, evaluate anticipated results, calculate sample sizes, and examine design considerations such as expected mean squares (see Engeman 1982). We have observed several other attributes that enhance the performance of a telephone statistician. The consultant should have a good telephone voice and personality. If the statistician sounds pleasant, intelligent, and interested over the phone, even the most tedious consulting sessions will progress more smoothly and the investigator will feel more inclined to keep in touch about the project. The consultant should have a strong ear and sufficient mental endurance to be effective through long phone conversations (up to several hours). It is also important for the statistician to have a flexible work schedule to allow for a quick response at home base to new problems in the field and, if necessary, to be able to go to the field on short notice.

Experience serves as the only teacher for most statisticians consulting long distance. We feel, however, that some awareness of the problems in nonvisual consulting could be developed in the classroom. To give students a feel for consulting without visual contact, an exercise involving two students visually separated by office dividers could be conducted in front of a classroom. One student could be given an experimental design or similar statistical problem that he or she should have to explain to the other without face-to-face communication. The task for the second student would be to understand quickly and respond to the problem. This exercise would allow the participating students to experience some of the difficulties of consulting in a nonvisual medium. With the dividers positioned correctly, the rest of the class could observe and later discuss the experience. This exercise could also be used as job training for statisticians already faced with long-distance consulting. Perhaps in conjunction with a psychology department, more elaborate training in consulting without face-to-face communication could be developed, but this relatively simple exercise should serve to open students' minds about the difficulties of nonvisual consulting.

The recommendations and approaches to statistical consulting made by other authors (described in the first paragraph of this article) provide an optimal approach to in-house consulting with clients. By comparison, consultation at a distance is more severely constrained and the statistician must be more attentive and creative to provide effective consultation. Our experience indicates that through the considerations discussed in this article, effective long-distance consultation can be accomplished.

[Received June 1983. Revised April 1984.]

REFERENCES

- BISHOP, THOMAS, PETERSEN, BRUCE, and TRAGSER, DAVID (1982), "Another Look at the Statistician's Role in Experimental Planning and Design," *The American Statistician*, 36, 387-389.
- ENGEMAN, RICHARD M. (1982), "Expected Mean Squares," *The American Statistician*, 36, 194.

FINNEY, D.J. (1982), "The Questioning Statistician," *Statistics in Medicine*, 1, 5-13.
 HAHN, GERALD J. (1984), "Experimental Design in the Complex World," *Technometrics*, 26, 19-31.
 HUNTER, WILLIAM G. (1981), "The Practice of Statistics: The Real World Is an Idea Whose Time Has Come," *The American Statistician*, 35, 72-76.

MARQUARDT, DONALD W. (1979), "Statistical Consulting in Industry," *The American Statistician*, 33, 102-107.
 McCULLOCH, C.E., BOROTO, D.R., MEETER, D., POLLAND, R., and ZAHN, D. (1982), "A Holistic Approach to Training Statistical Consultants," in *Proceedings of the Section on Statistical Education*, American Statistical Association, 116-121.

An Elementary Introduction to Maximum Likelihood Estimation for Multinomial Models: Birch's Theorem and the Delta Method

CHRISTOPHER COX*

A fairly complete introduction to the large sample theory of parametric multinomial models, suitable for a second-year graduate course in categorical data analysis, can be based on Birch's theorem (1964) and the delta method (Bishop, Fienberg, and Holland 1975). I present an elementary derivation of a version of Birch's theorem using the implicit function theorem from advanced calculus, which allows the presentation to be relatively self-contained. The use of the delta method in deriving asymptotic distributions is illustrated by Rao's (1973) result on the distribution of standardized residuals, which complements the presentation in Bishop, Fienberg, and Holland. The asymptotic theory is illustrated by two examples.

KEY WORDS: Multinomial models; Maximum likelihood; Birch's theorem; Delta method; Asymptotic distributions.

1. INTRODUCTION

Nearly every student of statistics receives some exposure to the analysis of categorical data. Advanced work in this area, including the study of log-linear models, can be based on the development of parametric models for the multinomial distribution. Consider, for example, a two-way contingency table with R rows, C columns, and multinomial probabilities p_{ij} ($1 \leq i \leq R$ and $1 \leq j \leq C$) satisfying $\sum_{i,j} p_{ij} = 1$. The independence model for the two classifications may be parameterized as $\log_e p_{ij} = \mu + \alpha_i + \beta_j$ for $\sum \alpha_i = 0 = \sum \beta_j$. Departures from independence are specified by a set of $(R - 1)(C - 1)$ interaction parameters. Placing

various restrictions on these parameters determines different families of models, such as the multiplicative interaction models considered by Goodman (1981). For the general log-linear model, let X be a $T \times s$ matrix of rank $s < T - 1$ whose columns are linearly independent of $\mathbf{1}$, a T vector of 1's. The multinomial probabilities are given, for $i = 1, \dots, T$, by

$$p_i = \exp(\mathbf{x}_i' \boldsymbol{\theta}) / \sum_j \exp(\mathbf{x}_j' \boldsymbol{\theta}),$$

where the column vector \mathbf{x}_i denotes the i th row of X . As an example that is not a log-linear model, consider the following expressions for the frequencies of human blood groups O, A, B, and AB (Kempthorne 1969). The four multinomial probabilities are specified by two parameters:

$$[(1 - p - q)^2, 2p(1 - q) - p^2, 2q(1 - p) - q^2, 2pq],$$

with admissible parameter set $\{(p, q) | p, q > 0 \text{ and } p + q < 1\}$.

A standard approach to parameter estimation for such models is the method of maximum likelihood. For families of distributions satisfying appropriate regularity conditions, standard large sample results (Lehmann 1980) guarantee the existence of solutions to the likelihood equations that are consistent, efficient, and asymptotically normal. For the special case of multinomial models, these results can be easily established by using Birch's theorem (1964) and the delta method (Bishop, Fienberg, and Holland 1975; BFH). This approach is less elegant but more general than the coordinate-free methods developed so thoroughly by Haberman (1974) for log-linear models. Like other "coordinate" approaches, it is also less abstract and simpler and therefore more suitable for an introductory presentation. To some extent the method also reveals more clearly the essential structure of the problem, since all of the results follow from the asymptotic distribution of the sample proportions by successive applications of the delta method. At the same time, an illustration is provided of this extremely useful large sample result.

Essentially, what is being called Birch's theorem states that for suitably regular parametric multinomial models, a

*Christopher Cox is Assistant Professor of Biostatistics in the Division of Biostatistics, University of Rochester School of Medicine and Dentistry, Box 630, 601 Elmwood Avenue, Rochester, NY 14642. Work was supported by National Institute of Environmental Health Sciences Grant ES-01248, by National Institute on Drug Abuse Grant DA-00623, and in part by U.S. Department of Energy Contract DE-AC02-76EV03490 with the University of Rochester, Department of Radiation, Biology, and Biophysics. The author would like to thank Christy Chuang and Charles Odoroff for helpful discussions and the referees and editor for constructive comments.