Statistical and Mathematical Aspects of Pollution Problems

edited by JOHN W. PRATT
Graduate School of Business Administration
Harvard University
Boston, Massachusetts

Statistical and Mathematical Aspects of Pollution Problems

edited by *JOHN W. PRATT*
Graduate School of Business Administration
Harvard University
Boston, Massachusetts

MARCEL DEKKER, INC. New York 1974

MARCEL DEKKER, INC.

305 East 45th Street, New York, New York 10017

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 73-90771

ISBN: 0-8247-6132-4

Current printing (last digit):

10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

LIST OF CONTRIBUTORS

D. BADELLON, Institut National de Recherche Chimique Appliquée, Vert-le Petit, France

M. BENARIE, Institut National de Recherche Chimique Appliquée, Vert-le Petit, France

L. A. CLARENBURG, Rijnmond Authority, Rotterdam, The Netherlands

STANLEY V. DAWSON, Harvard University, Boston, Massachusetts

ROBERT M. ELASHOFF, University of California at San Francisco, San Francisco, California

WILLIAM FAIRLEY, Harvard University, Cambridge, Massachusetts

JERZY GANCZARCZYK, University of Toronto, Toronto, Canada

J. GILLIGAN, Clyde River Purification Board, Glasgow, Scotland

D. HARVERSON, Edinburgh Corporation, Edinburgh, Scotland

V. HASSELBLAD, National Environmental Research Center, Environmental Protection Agency, Research Triangle Park, North Carolina

DAVID G. HOEL, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina

R. HUETING, Netherlands Central Bureau of Statistics, The Hague, Netherlands

ANDRÉ JUNOD, Institut Suisse de Météorologie, Payerne, Switzerland

ALBERT J. KLEE, Solid and Hazardous Waste Research Laboratory, Environmental Protection Agency, National Environmental Research Center, Cincinnati, Ohio

H. D. LANDAHL, University of California, San Francisco, California

LESTER B. LAVE, Carnegie-Mellon University, Pittsburgh, Pennsylvania

MICHAEL LEGRAND,*Institut Royal Météorologique de Belgique, Brussels, Belgium

G. R. LOWRIMORE, National Environmental Research Center, Environmental Protection Agency, Research Triangle Park, North Carolina

---

*Present address: Institut d'Hygiene et d'Epidemiologie, Brussels, Belgium.

D. W. MACKAY, Clyde River Purification Board, Glasgow, Scotland

CLIFFORD J. MALONEY, Food and Drug Administration, Bethesda, Maryland

T. MENARD, Institut National de Recherche Chimique Appliquée, Vert-le Petit, France

D. W. MOODY, U. S. Geological Survey, Washington, D. C.

W. C. NELSON, National Environmental Research Center, Environmental Protection Agency, Research Triangle Park, North Carolina

A. NONAT, Institut National de Recherche Chimique Appliquée, Vert-le Petit, France

R. P. OUELLETTE, The MITRE Corporation, McLean, Virginia

J. W. OVERBEY II, The MITRE Corporation, McLean, Virginia

ROBERT PIKUL, The MITRE Corporation, McLean, Virginia

M. A. SCHNEIDERMAN, National Cancer Institute, Bethesda, Maryland

EUGENE P. SESKIN, The Urban Institute, Washington, D. C.

MILTON SOBEL, University of California at San Francisco, San Francisco, California

DAVID STARRETT,* Harvard University, Cambridge, Massachusetts

MURRAY STEIN, Office of Water Enforcement, Environmental Protection Agency, Washington, D. C.

ROLF SUNDBERG,† Stockholm University, Stockholm, Sweden

SÖREN SVENSSON,‡ SCOPE Commission on Monitoring, Stockholm, Sweden

ROBERT V. THOMANN, Manhattan College, Bronx, New York

GERALD VAN BELLE, Florida State University, Tallahassee, Florida

J. R. WEST,** University of Dundee, Dundee, Scotland

D. J. A. WILLIAMS, University College, Swansea, Wales

THOMAS C. WINTER, JR., Council on Environmental Quality, Washington, D. C.

RICHARD ZECKHAUSER, Harvard University, Cambridge, Massachusetts

---

*Present address: Stanford University, Stanford, California.

†Present address: Royal Institute of Technology, Stockholm, Sweden.

‡Present address: Zoological Institute, Lund, Sweden.

**Present address: The University of Birmingham, Birmingham, England.

PREFACE

Pollution has become and will remain a problem of great concern to government, business, the public--and scientists. Statistical methods should play an important part in pollution studies and research, and conversely, playing its part should enrich the field of statistics in both methodology and applied experience. The papers in this book report on work being done of a statistical type, in the widest sense, on a number of kinds of pollution problems, and are representative of the best such work, insofar as the editor was able to identify it. They were written for a heterogeneous scientific audience and can be largely understood without specialized knowledge. As well as being of interest in themselves, then, they afford a broadly understandable survey of the state of the field, to the extent that this is possible in a book of such modest size. It has also proved possible, without impeding these uses of the book, to arrange the papers so as to provide an introduction to the application of statistical methodology in the context of pollution problems--not to each important statistical technique, but to each phase of the statistical method and some of the practical questions which arise therein, from the data collection program to the policy implications of the analysis. This will be explained more fully later.

The book originated in one of the Satellite Symposia sponsored by the International Association for Statistics in Physical Sciences (IASPS), an autonomous section of the International Statistical Institute (ISI), in connection with the 38th Session of the ISI which was held in Washington, D. C., August 10-20, 1971. The idea of the IASPS Satellite Symposia is well described in an article by P. A. P. Moran and J. Neyman, * who say that it was borrowed from the astronomers, although they attribute the name, oddly enough, to a statistical colleague. In brief, the purpose of the symposia was to foster cooperation between statisticians and substantive scientists on problems calling for skills of both types. A few sentences of their article are particularly relevant here:

> Of the interdisciplinary studies none seems as important as those relating to two international problems now widely discussed

---

* The IASPS and its Satellite Symposia, in The American Statistician, Feb. 1971, pp. 15-18, and the ISI Review, April, 1971.

> in scientific literature and in the daily press, the problem of the population explosion and the problem of environmental pollution. Both are interdisciplinary to an unprecedented degree and both relate to situations "at large" in which the establishment of pertinent facts must depend on sampling surveys, on randomized experiments, etc., that is, on the use of statistical methodology. The perusal of literature, often sensational and nearly always controversial, suggests skepticism as to the universal reliability of the underlying statistical studies, often performed by nonspecialists. On the other hand, the mass of important physical as well as biological knowledge involved in the two problems are overwhelming to statisticians.

Professor Neyman, as President of IASPS, asked me to organize a Satellite Symposium on Statistical Aspects of Pollution Problems. With the primary support of the National Science Foundation, it was held at the Harvard Business School, August 31-September 3, 1971. I believe that it served its purpose well. The participants were primarily responsible for this, of course, but, as an organizational point, it seemed to be facilitated, especially the important aspect of informal contact and discussion, by the location, including housing the participants together (in modest comfort but not hotel style), at a slight remove from the physical and mental pollution of Harvard Square.

The symposium was entirely invitational. I wanted papers by people doing real work of quality and value, viewed by informed scientists as leaders in such work, without regard to public notoriety or even to purely disciplinary reputation. I was in no position to make a list of such people myself, and I found that, in a field involving so many disciplines, it is difficult to identify them even second-hand (which perhaps adds support to the idea of a Satellite Symposium). I consulted too many people to name here but their scope is indicated by two who particularly deserve mention for help and helpfulness in getting me started: Myron B. Fiering, of the Division of Engineering and Applied Physics at Harvard, and David G. Kendall, a specialist in probability theory at Cambridge, England.

There was no advance commitment by anyone to any form of proceedings volume, but it was decided to undertake publication of all papers given at the symposium which were available within a reasonable period and were not being published elsewhere. A few of the written papers differed from the spoken ones. None were "refereed," but two or three have been completely or substantially rewritten by their authors since receipt, and two or three have been considerably cut. The rest have been edited, mostly slightly, a few more extensively, by the authors and/or editor, without any intention to change the substance appreciably. In the rare cases where it matters, the papers should be understood to apply as of the date of delivery. The discussion was excellent, but it seemed impractical to attempt to publish it, and only a small--but cogent--fraction appears here, essentially unedited. The following papers given at the symposium are not included in this book:

A mathematical model for pest control, Samprit Chatterjee, New York University; appearing in Biometrics, Dec. 1973.

Information-statistical evaluation of air pollution data samples, Camilo P. de La Riva, Landesanstalt für Immissions- und Bodennutzungsschutz des Landes Nordrhein-Westfalen.

A mathematical model for relating air quality measurements to air quality standards, Ralph I. Larsen, Environmental Protection Agency; published as Environmental Protection Agency Publication AP-89.

Invariant imbedding, nonlinear estimation, and the adaptive forecasting of stream pollution, E. Stanley Lee, Kansas State University.

Instabilities of regression estimates relating air pollution to mortality, Gary C. McDonald and Richard C. Schwing, General Motors Technical Center; appearing in Techonometrics, 15 (3), 463 (1973).

Advantages and limitations in storing, retrieval, and statistical evaluation of water quality data, Wolfgang Schmitz, Landesstelle für Gewässerkunde und Wasserwirtschaftliche Planung, Karlsruhe.

Evaluation of the APCO air quality display model calibration procedure, Max A. Woodbury, Duke University.

One area which would probably have been more heavily represented at the symposium but for conflicting meetings is biology and health. This gap was more than filled by a symposium in which Professor Neyman had a considerable direct part and which placed considerably more emphasis on airing controversies, with attendant advantages and disadvantages. Read all about it in the Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. VI, Effects of Pollution on Health (Lucien M. LeCam, Jerzy Neyman, and Elizabeth L. Scott, eds.), University of California Press, 1972. A critical survey of considerable literature by two participants in the IASPS symposium appears so useful that it deserves mention here, "Some Statistical Aspects of Environmental Pollution and Protection" by Gerald van Belle and Marvin Schneiderman, Florida State University Statistics Report M245, Nov. 1972, to appear in the ISI Review.

There are various ways to organize the field of pollution. A common one, which is useful for many purposes, is by medium: air, water, and soil, sometimes with additional categories such as biota, noise, and radiation. Another, obviously crosscutting, is by activity: power generation, transportation, etc. Stanley Dawson, one of the participants, separated the symposium papers neatly into four groups entitled Biological effects of air pollution and ingested substances, Measurement and analysis of air pollution, Waste disposal problems, and Strategies for control of environment. I chose a much more methodological organization for this book, however, for several reasons.

Much of the work described here, though without details if they are too complicated, is approaching the guts of a problem, and the real value of such work is often difficult to assess in isolation. The value of a data collection program depends on the inferences to be drawn from it and the effect of these inferences on policies and actions later implemented. The value of a mathematical model of the flow of a river depends on its use in this process. And so on. And the relative strength of much of the more specific work represented here is that it has a real use in this process. This is not to say that each paper considers everything from data collection to final decisions. That is ordinarily impossible to do in a modest space without becoming overly superficial. In fact, most of the papers rather clearly concentrate on a single aspect. But the ultimate goal is the improvement of the whole process of using data to reach decisions, rather than elaborating or refining one aspect for its own sake.

Organizing the book in accordance with this process therefore divides the papers naturally while emphasizing the importance of the whole process and clarifying the real purpose of the papers by relating them to it. More specifically, the process I mean is the now well-known statistical method (in whose elucidation, incidentally, Professor Neyman's role seems more and more to me to have been crucial). Each step will be discussed further later, but in brief: we have or obtain observations; we have or develop a model describing the probabilistic behavior of the observable data under various assumptions about the true state of the world; using the model and the observations, we make statistical inferences about the true state of the world; and, finally, decisions are made based on our inferences (we hope) and on other considerations of value, cost, feasibility, etc. The change to the passive voice accords with the fact that we (most of the authors, and I, if not the reader) make analyses as part of our profession, but will not make the final decisions. Despite this, however, the ultimate purpose of most of the work here is to aid decisions, not to advance pure science.

Reality may be much more complicated than this description, of course. In particular, we may recycle, from inference to further observation, for instance. And in designing an observational program or experiment, we should take at least informal account of all the subsequent phases of the process. But while considering them simultaneously we should not confuse them with one another.

The sections of the book relate to the process just described as follows. Parts I and II concern the collection and dissemination of basic data, and the general background and policy perspective to which decisions about the collection of data should be related. Part III concerns the development of models. Parts IV and V concern the drawing of statistical inferences from data using models.

To represent the remaining phase, I would have liked to have papers analyzing what decisions ought to be made, or what effects various decisions might have, on the basis of statistical inferences and other considerations including costs and the values that society might place on the possible outcomes. The absence of such papers at the symposium may represent bad luck or oversight. I suspect, however, that there is a dearth of such work, at least of a kind which takes explicit and analytical account of uncertainty. Likely reasons include the complexity of the decision problems and the diffused and political nature of decision making in this area. Also perhaps that decisions are not usually made until there is enough data to remove essentially all uncertainty (implying excessive delay or overspending on data). The mind's difficulty in encompassing uncertainty and the psychological tendency to block it may contribute. Unfortunately, the statistical concepts of "testing hypotheses" and "point estimation" may encourage this failing. Inferences incorporating a measure of uncertainty do not. Remember "confidence intervals"! Dare I suggest "posterior distributions"? I believe, and not just rhetorically, that both statistics and practice would gain if statisticians were more involved with this decision-making phase.

Comments follow on each Part in turn.

I. It seems natural to start with data--what data do you collect where, by what means, and how do you disseminate it or make it available? Part I consists of papers where these concerns dominate. They naturally emphasize large-scale, continuing programs and systems, but the possible efficiency of more specialized data-gathering efforts should not be forgotten. The most important question with large efforts is probably whether they should be undertaken at all. Admittedly my reading is far from comprehensive, but I cannot recall a single quantitative study of whether the gains from such an effort are likely to be worth the resources used, in pollution or any other field. A convincing job would be hard to do, but the only advantage of the rules of thumb one sometimes sees, stating an arbitrary percentage of the budget for data, is that they are cheap to compute. They certainly contradict what is known about optimal sample size in simple situations.

The first paper, by Svensson, sketches the thinking involved in the report of a high-level committee on environmental monitoring. This serves both to introduce the subject of pollution in its full international scope and to put forward some recommendations on data collection. These seem straightforward at first glance, but Clarenburg's bristling discussion promptly joins the issues, rapidly raising a host of difficulties.

The following two papers, by Junod and Legrand, illustrate local programs to collect air quality data, in areas of Switzerland and Belgium, Junod dealing particularly with the siting of the measuring stations and Legrand with the presentation of the data. (A more formal approach to the siting problem seems feasible, and likely to yield different rules from intuition.) Then Overbey and Ouellette discuss a data bank designed to collect centrally all such U. S. data, vast and disparate though it is, and to make it available in usable form. (They also illustrate its use, but if such illustrations appear to predominate, my editorial requests and inefficiency must be blamed for some distortion, as their real purpose was to describe the data bank.)

Other papers in the book also illustrate the nature and use of air quality data, especially the first three of Part IV. Some papers illustrating water quality data are the other two in Part IV and that by West and Williams (No. 13). All these later papers are, however, primarily concerned with analysis, rather than collection or dissemination per se. One other paper worth specific mention for relevance to Part I is that by Ganczarczyk (No. 22) who, in a very different situation (control of a sewage treatment plant), makes some pointed and possibly generalizable comments about data automation, quality, and usefulness.

II. A theoretically complete analysis of the design of an experiment or observational program involves considering, for each possible design, the model to be used in the analysis, the data which might be observed, and the resulting inferences, actions, costs, benefits, etc.--in short, the entire process through the final decision-making phase. And this must be done a priori, so that various possible outcomes must be considered with only subjective opinions available, not a posteriori when a particular outcome has occurred and evidence dominates opinion. More simply, how can you design an experiment without taking its analysis into account? Yet elementary courses in experimental design usually precede those in analysis, and one can go rather far in practical experimental design on the basis of informed common sense and rough notions of the decision problem. This may be true of purely observational programs too. Actually, what common sense handles fairly well is allocation of a given budget. It has little that is convincing to offer in setting the size of the budget, the total effort to expend in observation. Unfortunately, as mentioned earlier in a more limited way, analysts have not yet tackled this question much in complex, practical situations. One hopes the day will come when they are ready to, and a later day when funders are ready to understand the results. For now, however, the decisions will be made on the basis of general background and policy.

For this and other reasons, the papers in this book with predominantly a policy perspective are related to Part I (data) rather than to Parts III-V (probabilistic modeling and inference). Part II contains these papers, and those concerned with indices, inasmuch as indices are not basic data, but are combinations whose definition may reflect policy, and are an important form in which policy makers absorb information.

Opening Part II, Starrett and Zeckhauser approach a policy question (pollution control by markets or taxes ?) with economists' concepts and (deterministic) models. Though traditional assumptions and results fail, they explain and apply the concepts and way of thinking in a clear, elementary, and entertaining way in the context of pollution. Their consideration of economic efficiency may be juxtaposed with the consideration of enforceability in the following two papers by Stein and Winter, who represent the point of view of governmental policy (in the U. S.). Stein is concerned with water, while Winter discusses most other major topics. Next Pikul describes an effort under way to develop environmental indices. A frightening number may be required. Finally, Hueting presents a conceptual structure for the competition taking place in the use of the components of the environment for various purposes, giving particular attention to indices and to economic ideas.

III. The remaining papers in the book almost all concern the analysis of quite specific types of situation, and often of particular data. In Part III the dominant theme is the development of models, while in Parts IV and V it is the later stage of statistical analysis.

The first paper, by Clarenburg, illustrates excellently the complete process of data collection, probabilistic model building, statistical inference, and application to decisions. Particular steps can be questioned, but the process as a whole is very clear. Furthermore, of the individual aspects, models seem to me furthest developed. Thus the paper simultaneously serves well to introduce the whole body of analytical papers and fits into Part III. Van Belle's discussion, a fine example of reading both on and between the lines, provides much additional insight into what was actually done and raises many appropriate questions about it.

Clarenburg's specific subject is urban air pollution. The next two papers, also concerned with air pollution, touch two extremes of scale:

reduction in growth of forest trees (Sundberg) and deposition of particles in the human respiratory system (Landahl). The remaining two papers of Part III develop models for water movement. West and Williams treat an estuary one dimensionally (the Tay), while Harverson treats a bay two dimensionally (the Firth of Forth). For applications of standard statistical models in air pollution, see the first three papers of Part IV and those by van Belle and Maloney (Nos. 21 and 23); in water pollution, the final two papers of Part IV.

IV. As an alternative to building a special model for the problem at hand, it may be possible to use standard statistical models, some of which are very flexible. This offers obvious advantages, including well worked-out theory and, often, computer programs. The commonest statistical technique is simple cross-tabulation, where statistical inference beyond using observed proportions as estimates is often unnecessary and even more often omitted (except occasional ritual significance testing). Of the more advanced models and techniques, although some exist for contingency tables, regression is by far the most widely applied. Part IV contains five papers making heavy use of regression and related methods, other statistical methods being postponed to Part V. Both these parts naturally emphasize inference rather than model building, since prefabricated models are used.

Regression is such a powerful technique that it blinds us to its limitations. Deficient vocabulary compounds the problem. An increase of 1 unit in x is less easily said to "be accompanied by an increase in y of" beta units on the average than to "increase y by" this amount. Even such a correct phrase as "variance accounted for by x" is all too easily taken to mean "variance due to x," and how many can really remember in the heat of fray that the fraction of the variance of y accounted for by x coincides with the fraction of the variance of x accounted for by y? Or that the partial correlation of y with x_1 given x_2, x_3, ... equals that of x_1 with y? And when, in the physical nature of things, y cannot contribute causally to x but x can to y, even if the words speak only of association, it is hard to avoid attributing the association to the effect of x on y and to remember that the coefficient of x in a regression of y on x reflects not only any effect x has but also part of the effect of every omitted variable which is correlated with x.

To be explicit, with apologies for incomplete explanation, here is an approach I have found helpful. Suppose that the true causal model of y is

$$E(y \mid x, z) = \beta_0 + \beta_1 x + \beta_2 z,$$

so that β_1 measures the true effect of x on y, in which we are interested, and the only other variable having a true effect on y is z. Suppose that z is not observed, but its association with x, for whatever reason, leads to

$$E(z|x) = \alpha_0 + \alpha_1 x.$$

Then taking the conditional expectation of the first equation given x gives

$$E(y|x) = (\beta_0 + \beta_2\alpha_0) + (\beta_1 + \beta_2\alpha_1)x.$$

Thus in a regression model including x and excluding z, the coefficient of the included variable x differs from its true effect β_1 unless the excluded variable has either no effect or no association with x; the difference is the product of the effect of z on y and the coefficient of x in the regression of z on x. If x and z have more than one component, then β_2 becomes a row vector and α_1 a matrix, and the coefficient of an included variable represents its true effect if it has no association with any excluded variable having a true effect, given the other included variables. This can hold otherwise only if nonzero terms happen to cancel exactly.

The size of the effect of omitted variables can sometimes be bounded by plausible substantive arguments, with the help of the foregoing or some other technical approach. A model derived on the basis of theory has more causal plausibility than a purely observational relationship (which may push one toward custom-built rather than prefabricated models). And causation is supported if the coefficient of x is the same in different situations or populations (where the relation of z and x, measured by α_1 above, is presumably different) or when further variables are added (which has a similar effect). How much sameness is supported if it is accepted by a significance test is far from obvious, however, though it obviously depends on the power of the test to detect differences. High R^2 is not definitive: this can occur with cause and effect reversed or, more insidiously, through "spurious" correlations and "proxy" variables, especially for subsidiary causes.

Of course, the only sure way to measure causation is by experimentation, with controlled x's and appropriate randomization (making x ipso facto independent of all omitted variables). Unfortunately this road is often closed, and the alternative is arduous and hazardous. This is not to say regression should be avoided or ignored in nonexperimental situations. Other methods pose the same problems, perhaps hiding them even more, and serious analysis is always better than blind intuition or biased opinion. Nor is it to say, despite habits of thought derived from criminal law and statistical hypothesis testing, that it is always best to assume lack of causation until it is proved. Indeed, if this point is overlooked, it may sometimes lead to better ultimate decisions to overlook also the point that association does not prove causation. A little learning is a dangerous thing.

Lave and Seskin, in the first paper of Part IV, in connection with air pollution and mortality, provide a clear introduction to the regression model and the issues of association-causation-proof, though inevitably a misleading phrase or two must have escaped them (and me) which the sharp-eyed reader may find. Fairley's discussion carries the issues further.

In pure forecasting, the main purpose of the next paper by Benarie, Badellon, Menard, and Nonat, the usefulness of regression does not depend on causation, which therefore becomes peripheral or irrelevant. The inference to be drawn when large forecast errors occur raises similar issues, however, and Benarie et al. point out how, in their problem, the pattern of these errors affects the likelihood that the underlying situation has changed. Incidentally, it seems to me inappropriate to describe forecasting success by the significance level of a test of the null hypothesis that forecast and actual values are independent, because no one would be <u>that</u> pessimistic and because the significance level becomes extreme as more data is collected with no change in forecasting method or accuracy. Nevertheless, in correspondence Benarie informed me that in meteorology, although there is a lot of argument about how to do it, this is a common method.

In a third paper concerned with air pollution, Hasselblad, Nelson, and Lowrimore exemplify or touch on several extensions and relatives of regression. The remaining two papers of Part IV apply regression methods to water flow. Mackay and Gilligan study the Clyde estuary, while Thomann studies the Delaware--and demonstrates finally that the Scotch have no monopoly on water pollution analysis, at least. Thomann also demonstrates spectral analysis.

V. Papers predominantly concerned with statistical methods other than regression comprise Part V. In the first three, Klee describes a survey of attitudes on solid waste disposal, van Belle provides a one-man survey of statistical problems in aerosol studies, and Ganczarczyk discusses a number of aspects of activated sludge treatment plants.

The next three papers deal with three different kinds of problems involving dichotomous (binary, two-valued) data. Maloney introduces classical bioassay and indicates a special type of use it might have in pollution research. (Though he does not go on to such matters, analogues of the usual methods for continuous data have now been developed so far both theoretically and computationally that the conceptual unification which has long been possible should no longer be blocked by practical difficulties. An excellent reference by a leader in this work is Analysis of Binary Data by D. R. Cox, Methuen, 1970.) Elashoff, Sobel, and Schneiderman discuss the advantages and problems of designing screening experiments in which chemicals are given two instead of one at a time. (I conjecture that a Bayesian approach to this problem would be illuminating and not equivalent.) And Hoel, also concerned with testing chemicals, demonstrates a simple and surprising way in which a harmful chemical can appear beneficial and conversely.

Finally, for the last word, I have chosen some bittersweet comments by van Belle on responsible reporting and the role of statisticians in pollution research. Some of my own thoughts on other aspects of the latter, and on other topics, are in "Statistical Aspects of Pollution Problems, ISI Proceedings 44 (1971), Book 1, 181-6. Whatever I might have said here, I would certainly not have said as well as van Belle, so I will not try to top him.

Kyoto, Japan

John W. Pratt

ACKNOWLEDGMENTS

The National Science Foundation under Grant GP 30666 provided the main financial support for both the symposium and this book. Some additional and encouraging support was received from the National Bureau of Standards. The travel expenses of many participants, especially employees of U. S. government agencies, were paid for by their employers or other sources of funds, an important factor in paring the budget. My time in editing has been partly supported by sabbatical leave from Harvard and by the John Simon Guggenheim Memorial Foundation, with excellent working facilities provided by the Kyoto Institute of Economic Research of Kyoto University. All these sources of support are gratefully acknowledged.

David Lynn helped with many aspects of putting the program together. Sandra MacArthur, my secretary at the time, performed nobly a variety of tasks extending far beyond ordinary secretarial ones. Edward D. Rowley took unstinting care of local arrangements and preoccupied conferees. Among them, they took everything off my hands I could hope for, and I am most grateful to them.

Finally, although I have not sought the authority of the authors to dedicate this book formally, I would like to dedicate at least my part in it to Jerzy Neyman, in recognition of his leading role in bringing statistics to its present stage of maturity and in conceiving the IASPS and its Satellite Symposia and bringing them into being by force of intellect, diplomacy, persistence, and sheer will.

John W. Pratt