sentation of points in closed bounded convex separable subspaces of spaces with the RNP and Phelp's theorem that X possesses the RNP if and only if each nonempty closed bounded convex set in X is the closed convex hull of its strongly exposed points.

In sum, the book is a valuable source to workers in the area of Banach spaces. It is full of details and proofs which are concisely and clearly presented. It is a welcome addition to the growing number of books on Banach spaces.

H. ELTON LACEY

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 5, September 1976

Statistical prediction analysis, by J. Aitchison and I. R. Dunsmore, Cambridge Univ. Press, London, New York, Melbourne, 1975, xi + 273 pp., $24.50.

No doubt future statisticians will find it remarkable that not before the last quarter of the 20th century, did a textbook on what Statistics is really about, finally appear. Too long has the estimation of parameters dominated statistical theory and consequently warped and cluttered up the methodology–the raison d'être of the field, while the prediction of observables, which should have been preeminent, receded into the background. There are several reasons why this occurred. Chief among them is the tremendous preoccupation that theoreticians have had analysing the logical distinctions inherent in the various so-called modes of Inference, i.e., Bayesian (Jefferies, de Finetti, Savage), Frequentist (Neyman-Pearson), Fiducial (Fisher), Likelihood (Fisher, Barnard), etc., rather than what should be the proper subject for statistical analysis–parameters or potential observables. In the early history of Statistics there was no sharp distinction drawn between statistics (functions of observables) and hypothetical parameters, resulting in the tendency for the issue to be obscured. R. A. Fisher correctly made the sharp distinction necessary for the advance of thinking in this area. But since then and apparently through no fault of Fisher's, mathematical statisticians became so enamoured of those artificial constructs–parameters, that all their work tended to be framed and executed parametrically. Oddly enough even in that branch of Statistics which is often referred to as Non-Parametric Inference, developers and practitioners also attempt, to this day, to orient their work towards the estimation of parameters–so ingrained is the habit. Some, perhaps realizing the paradox, even altered the taxonomy by referring to this branch as Distribution-Free Inference.

One must also realize that the parametric approach has advantages, though illusory. Mathematical statisticians using any mode were often seduced by the niceties of the mathematics of parametric structures. Making precise statements about unobservables, i.e., parameters, also serves applied statisticians very well in that it is virtually impossible to contradict them by observation. Of course a predictivist, who by definition is in the business of making statements about potential observables, lacks such security. His statements, to a large extent, can usually be assessed by further observations from the process in question or by withholding at random some of the observations already at hand and empirically evaluating the accuracy of his statements.

Since the early part of the last decade I was certain that the inferential stress on estimation was misplaced and that prediction should be preeminent but put forward my notions rather timidly mainly in the guise of producing predictive distributions, e.g., Geisser [**1964**], [**1965**], [**1966**], Enis and Geisser [**1971**]. However this culminated with a mildly polemical paper, Geisser [**1971**], where the notion, partially implemented, was put forth that most if not all of the so-called estimation problems could be handled within a predictivistic framework–and more informatively! For the interested reader, somewhat stronger polemics can be found subsequently in Geisser [**1973a**], [**1973b**], [**1974b**], [**1975**], [**1976**].

Basically a statistical prediction problem, my contention being that almost all statistical problems could be cast as such, is an incompletely specified random process which, even when completely specified, will generate observations that cannot be completely determined beforehand. The process, often an artificial construct, is usually characterized by parameters for convenience. At the risk of being redundant I will again state that statistical theory has focused on these intermediary nuisances to the neglect of the real problem–the distribution of future observables or some scaling of plausibilities for their possible values depending only on the assumptions and the data at hand.

All this background serves two purposes. It gives the reader a thumbnail sketch of where I believe modern statistics has gone off the track and indicates my great pleasure in seeing corrective action being taken where it will do the most good–at the textbook level.

With this in mind I note that the authors have immersed but a single foot into the predictive sea, however one of uncommon competence. While averring that prediction is at the heart of many statistical applications, they have mainly restricted their discourse to cases traditionally associated with prediction–exceptions being classification and some associated areas. Predictivism is then only partially embraced and consequently the challenge, Geisser [**1971**], of presenting all statistical analyses in the predictive mode is handled in a rather conservative fashion and in some respects ignored. Although this is regrettable it would be equally unseemly to chide them for not having the courage of the reviewer's convictions.

There is an introductory chapter setting down the assumptions relating a future data set to a current set and indicating the kinds of applications to be considered. The second chapter uses the Bayesian approach to calculate predictive distributions for future observations. It also contains a catalogue of common sampling distributions and predictive distributions obtained from prior distributions closed under sampling. At the end of this chapter and in each subsequent chapter there is a section entitled History which in a few cases, not entirely due to brevity, is misleading and revisionist in that priorities are presumed that are either incorrect or obscure and important papers neglected.

In the third chapter, predictive utilities are introduced, i.e., the utility of taking action a, on the basis of potential observation y, or a set of such values rather than on the assumption of the truth of parameter value θ–the customary utility function. The relation between the two utilities is carefully unraveled. The next chapter, entitled "Informative Prediction" discusses ways of using predictive distributions in problems of coverage tolerance and other areas where the assessment of the predictor is not strictly tied to a utility, or utilities are not easily specifiable. Up to this point the view is strictly Bayesian. The next two chapters switch inferential modes and describe tolerance prediction in frequential terms. Although the authors stress here that a predictor of "similar mean coverage" has the same "mean Bayesian coverage" irrespective of the prior density used, the usefulness of this result is obscure for a Bayesian. The averages are over joint repetitions of current and future data sets but the appropriate reference set should be restricted to future repetitions given the current data in hand. In other words the value of the frequentist mode is in pretrial assessment, but is not relevant once the sample is in hand. However one must concede that the authors have presented frequentist tolerance theory in as clear and lucid a form as possible. Actually as long as the authors were willing to present deviationist philosophy, they could have at least included relative likelihood, second order likelihood, and predictive likelihood approaches, e.g., Fisher [**1956**], Kalbfleisch [**1971**], Lauritzen [**1974**]. Here, without the full Bayesian apparatus, values for future observations are assigned measures of plausibility, conditional on past values. This seems to be more informative than what the usual frequentist approach yields.

In Chapter 7 there is a section on finding predictive distributions by assigning a prior distribution for the future observation rather than for the parameter, before taking data. This is, of course, the ultimate objective of predictivists, to attempt to bypass the whole parametric maze, though it is not entirely clear to what degree such a program is realizable. This is technically an old area of research termed identifiability, which now should receive new impetus because of its importance in predictivism. The rudiments of empirical Bayes prediction are also presented as well as distribution-free tolerance intervals utilizing order statistics. Unfortunately the newest form of low structure prediction developed independently by Stone [**1974**] and Geisser [**1974a**], [**1975**], based on sample reuse techniques which simulate prediction, are not included.

Sampling inspection involving decisions about items in a lot based on the testing of a fraction of those items with utilities defined for the untested items, is the subject of the next chapter. Again it is transparent, especially from the practical point of view, that this is a much more sensible way of handling the problem than using parametric utilities. This bears a strong resemblance to sample survey problems where one starts with a finite number of items, samples some and tries to infer what the rest are like which surely is the fundamental problem of Statistics. On p. 147 of the text the authors make a simplifying assumption that the utility is additive and identical for each of the untested $N - n$ observations. This allows the calculation of the total utility to be $N - n$ times the common utility for each untested observation. The demonstration here is longer than necessary. Since the joint predictive distribution of the untested items in this setup must yield identical marginal distributions (in fact they are technically exchangeable), the proof can proceed directly on the joint predictive distribution without referring back to distributions depending on parameters. Topics such as optimum fixed sample size choice and sequential schemes are also presented.

Chapter 9 deals with problems of regulation and optimization. Here a series of n experiments, indexed by $t \in T$, are performed with outcomes $\{(t_1, x_1), \ldots, (t_n, x_n)\}$. In the regulation problem we direct our future experiment towards producing a value y_0 or a value in some set Y_0 by choosing or controlling the index t in attaining this goal. In the optimization case the goal is generally to obtain an optimal value in some sense for a future experiment by suitable choice of t. This is one of the few areas where point prediction is entirely sensible since a single future experiment is to be conducted. It would appear that if one envisaged several possible future experiments for either regulation or optimization that a sequential predictive procedure would be most natural and the method could easily be adapted to such a contingency.

If instead of choosing a future experiment, it had already been performed and one's goal is to identify the index t which was unknown from the known outcome y, we have the problem of calibration, the subject of Chapter 10. Instead of this being a prediction problem it is a retrodiction problem as the classification problem often is. But this is temporal quibbling. As long as one is careful about the assumptions one can predict backwards as well as forwards. Here one obtains a predictive or, as the authors term it, a calibrative distribution of unknowns given known outcomes. Various possible assumptions for natural and designed calibrative experiments are carefully given in detail. I have one minor objection to the fact that the authors designate an observable or potentially observable entity as a parameter. I prefer to reserve the term parameter for those entities which are not physically capable of being observed. At any rate this would be the predictivistic view as I see it. This chapter ends with an insightful comparison of several classical calibration predictors with those generated from the Bayesian approach.

Although the authors give an undeserved prominence to medical problems by titling the 11th chapter, "Diagnosis", there is much to commend this designation in place of the more usual ones such as Discrimination, Classification or Allocation. Actually Diagnosis is less action oriented and emphasizes the retrodictive context of many applications. The problem being to ascertain the origin of an observation by means of its values on a variety of characteristics on the assumption it arose from one of several mutually exclusive and exhaustive categories or populations. Information is at hand in terms of samples from each of the populations which are specified up to some unknown parameter set. The predictive probability that the observation belongs to each of the several populations is computed via the Bayes approach.

The authors restrict this predictive allocation procedure to a single observation which could be termed marginal allocation. As was shown, Geisser [**1966**], this marginal procedure needs to be modified for optimal joint allocation of several new observations or for sequential allocation where a temporal imperative exists in that new cases must be "diagnosed" or allocated as soon as they arise. The predictive methods are compared with the older estimative methods which insert estimates for parameters in conditional probabilities with the result that the latter are found wanting. The last chapter titled "Treatment Allocation" employs the predictive approach to problems of optimal selection.

Throughout the book there are excellent tables of common distributions, a few given uncommon names and notational designations, but nevertheless very helpful to the reader. At the end there are two useful appendices, one relating to distributions and relations among them and the other to classification of the prediction problems considered in the text.

It is not unusual for a statistical textbook to be either a trite mathematical exposition of what purports to be statistical theory indistinguishable from an ordinary mathematics text or at the other extreme a catalogue of recipes embellished with more or less concocted illustrative examples. Not so here. The material is presented with great care and consummate skill and is larded with incisive analyses of real experimental data culled from the fields of medicine, finance, biochemistry, biology, industrial engineering and agriculture–a broad mix indicating the purview of the methodology. For two reasons this is a rare and important textbook. Important because it is the first to address itself to predictivism and rare due to its expository clarity and lucidity and both for the care and effort that went into the work. So much so the latter that the reviewer failed to uncover a noticeable misprint or error. No doubt there are some, but they must be few and far between.

REFERENCES

P. Enis and S. Geisser [**1971**], *Estimation of the probability that $Y < X$*, J. Amer. Statist. Assoc. **66**, 162–168. MR **43** #4179.

R. A. Fisher [**1959**], *Statistical methods and scientific inference*, 2nd rev. ed., Hafner, New York. MR **24** #A1756.

S. Geisser [**1964**], *Posterior odds for multivariate normal classifications*, J. Roy. Statist. Soc. Ser. B, pp. 69–76. MR **30** #4340.

——— [**1965**], *Bayesian estimation in multivariate analysis*, Ann. Math. Statistic. **36**, 150–159. MR **30** #4341.

——— [**1966**], *Predictive discrimination*, Multivariate Analysis (Proc. Internat. Sympos, Dayton, Ohio, 1965; P. R. Krishnaiah, Ed.), Academic Press, New York, pp. 149–163. MR **35** #2417.

——— [**1971**], *The inferential use of predictive distributions*, Foundations of Statistical Inference (V. Godambe and D. Sprott, Eds.), Holt, Rinehart and Winston, Toronto, pp. 456–469.

——— [**1973a**], *Discussion on "Marginalization paradoxes in Bayesian and structural inference by Dawid et al"*, J. of Roy. Statist. Soc. Ser. B **35** (2), pp. 224–225.

——— [**1973b**], *Discussion on "Bayesianism: Its unifying role for both the foundations and the applications of Statistics by de Finetti"* (Proc. 39th Internat. Statist. Inst., Vienna, 1973), Bull. Inst. Internat. Statist **45**, book 4, p. 374. MR **50** #15022.

——— [**1974a**], *A predictive approach to the random effect model*, Biometrika, **61** no. 1, pp. 101–107.

S. Geisser [1974b], *Discussion on "Cross-validatory choice and assessment of statistical predictions by Stone"*, J. Roy. Statist. Soc. Ser. B, **36** (2), pp. 141–142. MR **50** #8847.

——— [**1975**], *The predictive sample reuse method with applications*, J. Amer. Statist. Assoc. **70**, no. 350, pp. 320–328.

——— [**1976**], *Predictivism and sample reuse*, 21st Design of Experiments Conf. (in press).

J. D. Kalbfleisch [**1971**], *Likelihood methods of prediction*, Foundations of Statistical Inference (V. Godambe and D. Sprott, Eds.) Holt, Rinehart and Winston, Toronto, pp. 378–392.

S. L. Lauritzen [**1974**], *Sufficiency, prediction and extreme models*, Scand. Stat., pp. 128–134.

M. Stone [**1974**], *Cross-validatory choice and assessment of statistical predictions*, J. Roy. Statist. Soc. Ser. B, **36** (2), pp. 111–147. MR **50** #8847.

SEYMOUR GEISSER

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 5, September 1976

$SL_2(\mathbf{R})$, by Serge Lang, Addison-Wesley Publishing Co., Reading, Massachusetts, 1975, 425 + xvi pp., \$19.50.

Given the formalism of quantum mechanics, the study of those of its laws which are invariant under the Lorentz group inevitably leads to infinite-dimensional representations of both the homogeneous and inhomogeneous forms of the group. Responding to earlier work of Dirac and Wigner, Bargmann, Gelfand-Naĭmark, and Harish-Chandra published in 1946 and 1947 classifications of the unitary representations of the homogeneous Lorentz group, or rather of a covering group, SL(2, **C**). Bargmann, because he was interested in the representations of the inhomogeneous group, was led as well to classify the representations of SL(2, **R**), a covering group of the Lorentz group in three variables, two in space and one in time.

From these innocent beginnings the mathematical theory of infinite-dimensional representations has expanded relentlessly, forgetting its origins in physics but encroaching on other domains of mathematics, especially number theory, to which its methods, those of functional analysis with a heavy admixture of Lie theory, have been foreign. Since the training of many contemporary number-theorists has been primarily algebraic, even those who view the new methods with favor find them difficult to assimilate. Some simple introductions are needed, not so much to expose the techniques, or even the basic concepts, but just to pierce the tough rind of unfamiliarity. Such is the purpose of $SL_2(\mathbf{R})$. It is a rough-hewn book, leisurely and informal, which in the manner of a good graduate course, conscientiously explains the heterogeneous facts from various domains which could be stumbling blocks for the novice, and may be exactly what is needed.

Bargmann, to whose paper many later students have turned for an introduction to the subject, discovered in particular a discrete series of irreducible representations of SL(2, **R**) with square-integrable matrix coefficients. It is remarkable that most of the phenomena which are significant for the general theory, not only the discrete series, whose importance cannot be exaggerated, but also other things more easily overlooked, appear already in SL(2, **R**); those who are led through it by an experienced guide will, if they later penetrate the general theory, meet nothing totally unfamiliar.

But one does not reach new continents by skirting the coasts of home. The physicist is concerned almost exclusively with the internal structure of the